ICMR-CDSCO STANDARD PERFORMANCE EVALUATION PROTOCOLS

DIVISION OF COMMUNICABLE DISEASES, ICMR
IN VITRO DIAGNOSTICS DIVISION, CDSCO

SEPTEMBER, 2025

New Delhi, India

सत्यमेव जयते

डॉ. राजीव बहल, एमडी, पीएचडी
DR. RAJIV BAHL, MD, PhD

INDIAN COUNCIL OF MEDICAL RESEARCH

सचिव, भारत सरकार
स्वास्थ्य अनुसंधान विभाग
स्वास्थ्य एवं परिवार कल्याण मंत्रालय एवं
महानिदेशक
भारतीय आयुर्विज्ञान अनुसंधान परिषद

Secretary, Government of India
Department of Health Research
Ministry of Health & Family Welfare &
Director-General
Indian Council of Medical Research

File No. ICMR-CDSCO/IVD/2023-Part(2)
Dated 4th September 2025

FOREWORD BY DIRECTOR-GENERAL, INDIAN COUNCIL OF MEDICAL RESEARCH

After the COVID-19 pandemic, the diagnostic industry of India has grown immensely, and is ready to contribute to the public health system by developing in-vitro diagnostic (IVD) tests for timely detection of pathogens causing critical infectious diseases. It is important to keep pace with the enthusiasm of our diagnostic industry and create commensurate mechanisms for technical oversight, such as standard in-vitro diagnostics evaluation protocols, and a network of accredited laboratories for timely validation of new diagnostic kits and their introduction in the market.

The Indian Council of Medical Research (ICMR) and the Central Drugs Standard Control Organization (CDSCO) are working tirelessly to create this ecosystem within the country to support in-vitro diagnostic kit validation. I am happy to note that the first batch of 39 standard IVD evaluation protocols are being unveiled. This includes protocols for performance evaluation of IVDs detecting Tuberculosis, Malaria, Dengue, Chikungunya, Zika virus, Typhoid fever, various respiratory viruses, Chandipura virus and Nipah virus. I am sure that this landmark initiative will contribute immensely in strengthening the quality of IVDs in India, and also enable them to compete for global licensure.

I congratulate the ICMR and CDSCO teams for undertaking this excellent endeavor. I also understand that the process of standard protocol development will continue, and the scope will be expanded to cover other diseases and IVD formats.

Rajiv Bahl

(Dr. Rajiv Bahl)

V. Ramalingaswamy Bhawan, Ansari Nagar, New Delhi-110029
Tel. : 011-26588204, Website : www.icmr.gov.in, E-mail : secy-dg@icmr.gov.in

डॉ.राजीव सिंह रघुवंशी
औषधि महानियंत्रक (भारत)
केंद्रीय औषधि मानक नियंत्रण संगठन
स्वास्थ्य एवम परिवार कल्याण मंत्रालय
भारत सरकार
एफ.डी.ए. भवन, कोटला रोड,
नई दिल्ली–110002

सत्यमेव जयते

Dr. Rajeev Singh Raghuvanshi
Drugs Controller General (India)
Central Drugs Standard Control Organisation
Ministry of Health & Family Welfare
Government of India
FDA Bhawan, Kotla Road
New Delhi-110002 (India)

08 SEP 2025

FOREWORD

In-vitro Diagnostic (IVDs) play a vital role in early diagnosis of disease, support surveillance and enhance patient care managements. As the healthcare landscape in India continues to evolve, ensuring the quality and performance of IVDs in pre and post market phase is crucial for accurate diagnosis and effective treatment.

It gives me immense pleasure to present this important compendium of Standard In-Vitro Diagnostic (IVD) Evaluation Protocols, jointly developed by the Indian Council of Medical Research (ICMR) and the Central Drugs Standard Control Organization (CDSCO). This collaborative effort marks a significant milestone of the MoU between ICMR and CDSCO aimed at strengthening the IVD testing capacity. This is the first initiative of its kind globally, where standardized evaluation protocols have been developed in collaboration between National regulatory authority and National nodal research Institute to assess the performance of In-Vitro Diagnostic (IVD) kits used for high risk diseases. This document will also serve as a valuable resource by providing a uniform testing protocol and setting minimum standards and acceptance criteria for various critical disease markers.

By streamlining the evaluation process, we also hope to accelerate the availability of high-quality diagnostic products to clinicians and patients, supporting the national vision of affordable, accessible and quality healthcare for all. We will continue to build on this initiative by adding more protocols and establishing more validation centers, ensuring that our regulatory framework remains robust and effective.

I congratulate the teams at Division of communicable disease, ICMR and IVD Division, CDSCO for their tireless efforts in drafting these protocols. I am confident that this publication will contribute significantly to the advancement of diagnostic science and regulatory excellence in India.

Dr. Rajeev Singh Raghuvanshi
Drugs Controller General of India

Tel.: +91-11-23236965 Fax : +91-11-23236973 Email : dci@nic.in Website : www.cdsco.gov.in

Contents

ARBOVIRUS IN-VITRO DIAGNOSTICS

List of Contributors:

A. Working Group:

1. Dr. Alagarasu Kalichamy, Scientist-E, ICMR-National Institute of Virology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
2. Dr. Megha Brijwal, Additional Professor, Department of Microbiology, All India Institute of Medical Sciences, Delhi
3. Dr. Gururaj Rao Deshpande, Scientist-C, ICMR-National Institute of Virology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
4. Dr. M. Santhosh Kumar, Scientist D, ICMR-National Institute of Epidemiology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
5. Ms. Krittika Bhattacharyya, Statistical Officer (Planning), Directorate of Economics and Statistics, Government of National Capital Territory of Delhi
6. Dr. Vishal Deo, Scientist-C, ICMR-ICMR- National Institute for Research in Digital Health and Data Science, Department of Health Research, Ministry of Health and Family Welfare, Government of India
7. Dr. Labanya Mukhopadhyay, Scientist-C, Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

B. Review Committee:

1. Dr. Vasanthapuram Ravi, Former Dean Research and Head of Neurovirology, National Institute of Mental Health and Neuro-Sciences, Bengaluru, Karnataka
2. Dr. Amita Jain, Former Professor and Head of the Department of Microbiology, King George's Medical University, Lucknow, Uttar Pradesh
3. Dr. Manoj Murhekar, Scientist-G and Director, ICMR-National Institute of Epidemiology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
4. Dr. Lalit Dar, Professor and Head of the Department of Microbiology, All India Institute of Medical Sciences, Delhi
5. Dr. Gajanan Sapkal, Scientist-F, ICMR-National Institute of Virology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
6. Dr. Guruprasad Medigeshi, Professor, Indian Institute of Science Education and Research, Tirupati, Andhra Pradesh
7. Mr. Pramod Meshram, Deputy Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
8. Dr. Sella Senthil, Assistant Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
9. Dr. Md Omair Anwar, Drugs Inspector (Medical Devices), IVD Division, Central Drugs Standard Control Organization, Government of India
10. Dr. Nivedita Gupta, Scientist-G and Head of the Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

Performance evaluation protocol for Chikungunya IgM ELISA kits

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured diagnostic kits appropriate for use in India. Hence the following guidelines shall establish uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Chikungunya IgM ELISA kits in the diagnosis of Chikungunya infection using irreversibly de-identified leftover archived/ spiked clinical samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through

A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

- Preparation & characterization of kit evaluation panel

- Handling of Chikungunya IgM ELISA kits received for performance evaluation (Verification/Storage/Unpacking etc).
- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Chikungunya IgM ELISA IVD kit evaluation panel:

Well characterised Chikungunya IVD kit evaluation panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of sera samples should be available from Chikungunya confirmed cases. Further characterised for Chikungunya IgM positivity by using approved reference kits having high sensitivity and specificity.

Chikungunya IgM performance evaluation panel need to be tested again by the reference assays at the time of evaluating a particular index test to confirm the positive and negative status of the samples.

4. Reference assay:

All the samples will be tested by CDC/NIV real-time (RT-PCR) assay. *Samples which are positive by RT-PCR assay will be further tested by any two of the following Chikungunya IgM ELISA kits:*

i. *ICMR-NIV MAC ELISA kit*
ii. *Inbios* CHIKjj Detect™ IgM ELISA
iii. Anti-Chikungunya virus ELISA (IgM) Test (Euroimmun, Luebeck, Germany)
iv. Any WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved IgM ELISA (as and when available)

Samples positive by PCR and at least two kits will be considered positive. If sufficient RT-PCR positive samples are not available, samples positive by at least 2 ELISA kits (of the kits mentioned above) can be considered as true positive samples.

Samples which are negative by RT-PCR and at least two IgM ELISA kits mentioned above will be considered as Chikungunya negative samples.

5. Sample size and sample panel composition*:* Sample sizes of positive and negative samples and sample panel composition against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, and an absolute precision of 5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\,(1 - S_e)}{d^2}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\,(1 - S_p)}{d^2}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*

Positive samples: Positive samples should be positive by RT-PCR at least two ELISA kits from the three mentioned above. If sufficient RT-PCR positive samples are not available, samples positive by at least 2 ELISA kits (of the kits mentioned above) can be considered as true positive samples.

Negative samples: Samples which are negative by RT-PCR and at least two IgM ELISA kits mentioned above will be considered as Chikungunya negative samples.

Table 1. Sample sizes and panel composition of positive chikungunya samples for different values of sensitivity claimed by the manufacturer

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%	15	20	Strong positive: 6 Moderate positive: 7 Weak positive: 7
95%	73	80	Strong positive: 24 Moderate positive: 28 Weak positive: 28
90%	138	140	Strong positive: 42 Moderate positive: 49 Weak positive: 49
85%	196	200	Strong positive: 60 Moderate positive: 70 Weak positive: 70

80%	246	250	Strong positive: 75 Moderate positive: 87 Weak positive: 88
The samples need to be classified as strong, moderate and weak positives based on ELISA units of the reference assay. *#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.*			

Table 2. Sample sizes and panel composition of negative chikungunya samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required [Sample size rounded off]#*	*Sample Panel Composition*
99%	15	20	1. Cross reactivity panel: • Dengue IgM positive: 3 @ • Rubella IgM positive: 1 * 2. [a]Acute febrile illness cases: 8 3. [b]Healthy subjects from endemic regions: 8
95%	73	80	1. Cross reactivity panel: • Dengue IgM positive: 15 @ • Rubella IgM positive: 5 * 2. [a]Acute febrile illness cases: 30 3. [b]Healthy subjects from endemic regions: 30
90%	138	140	1. Cross reactivity panel: • Dengue IgM positive: 26 @ • Rubella IgM positive: 8 * 2. [a]Acute febrile illness cases: 53 3. [b]Healthy subjects from endemic regions: 53
85%	196	200	1. Cross reactivity panel: • Dengue IgM positive: 38 @ • Rubella IgM positive: 12 * 2. [a]Acute febrile illness cases: 75 3. [b]Healthy subjects from endemic regions: 75
80%	246	250	1. Cross reactivity panel: • Dengue IgM positive: 47 @ • Rubella IgM positive: 15 * 2. [a]Acute febrile illness cases: 94 3. [b]Healthy subjects from endemic regions: 94
[a] Acute febrile illness cases negative for above pathogens AND Chikungunya IgM & PCR [b] Samples from healthy subjects from endemic regions negative for all Chikungunya markers (IgM, RNA)			

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

6. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

7. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

8. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

9. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

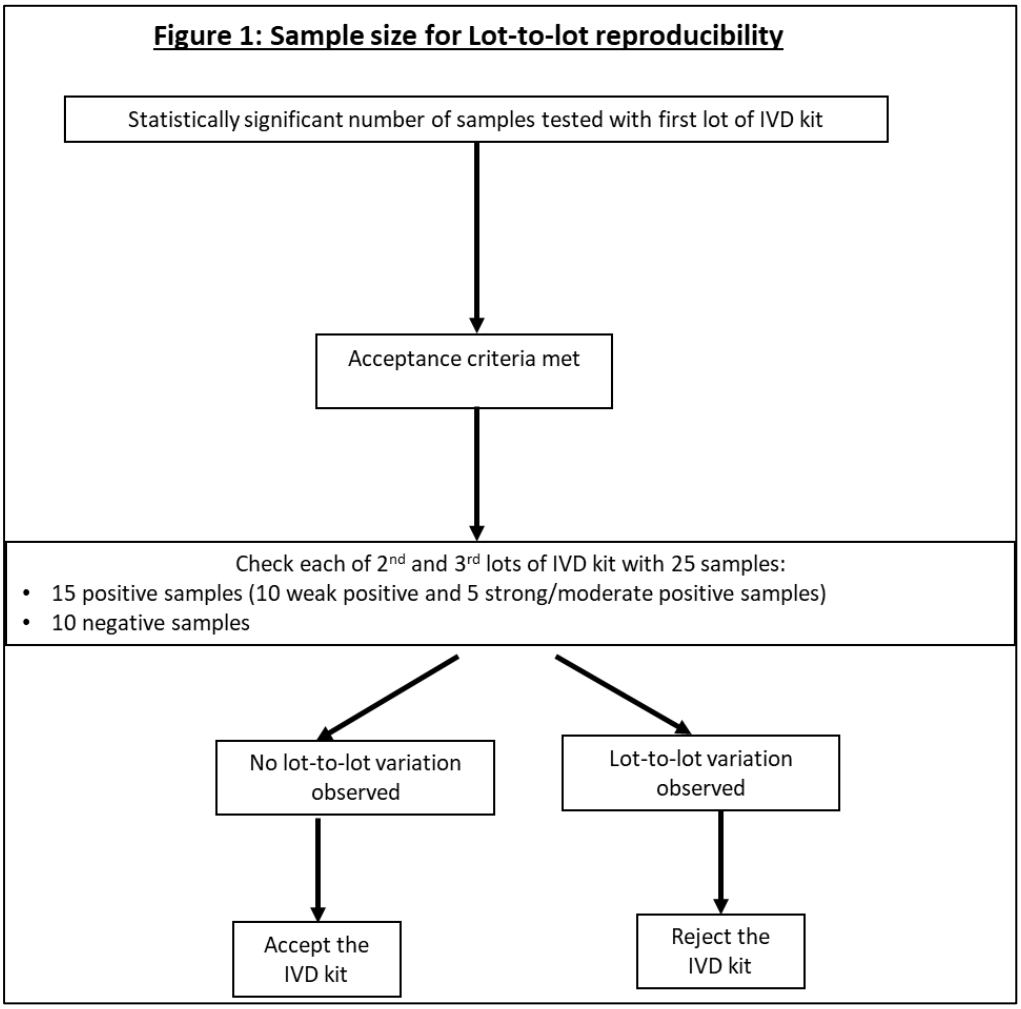
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

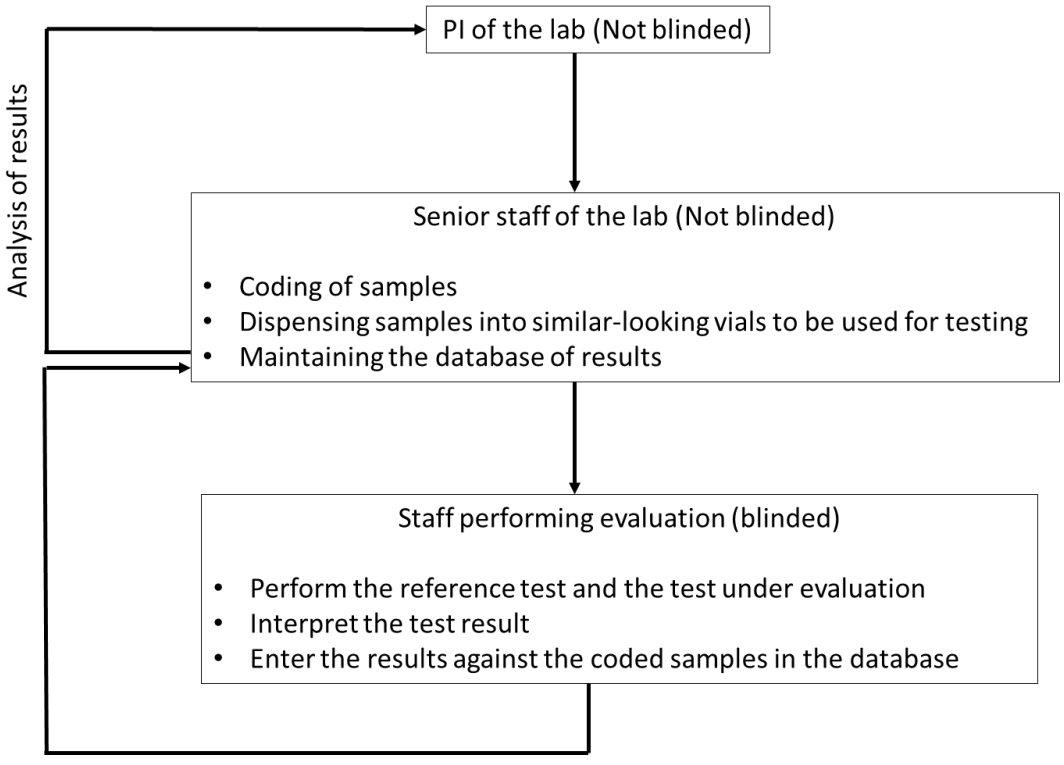
c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

10. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



11. Acceptance Criteria:

Sensitivity: ≥90%

Specificity: ≥95%

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥140 positive samples and ≥80 negative samples should be used for evaluation.

12. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Kikuti M, Tauro LB, Moreira PSS, et al. Evaluation of two commercially available Chikungunya virus IgM enzyme-linked immunoassays (ELISA) in a setting of concomitant transmission of Chikungunya, Dengue and Zika viruses. Int J Infect Dis. 2020 Feb;91:38-43.
2. Johnson BW, Goodman CH, Holloway K, de Salazar PM, Valadere AM, Drebot MA. Evaluation of Commercially Available Chikungunya Virus Immunoglobulin M Detection Assays. Am J Trop Med Hyg. 2016 Jul 6;95(1):182-192. doi: 10.4269/ajtmh.16-0013. Epub 2016 Mar 14.
3. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
4. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental, 8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR CHIKUNGUNYA IgM ELISA KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak/simulated samples)	
	Negative samples (provide detail: clinical/spiked, including cross reactivity panel/simulated samples)	

Results:

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Chikungunya IgM ELISA kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

Conclusions:

- Cross-reactivity:
- Invalid test rate:

- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ….. sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for ……Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……

Performance evaluation protocol for Chikungunya IgM RDT kits

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Chikungunya IgM RDT kits in the diagnosis of Chikungunya infection using irreversibly de-identified leftover archived/ spiked clinical samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through

   A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel

- Handling of Chikungunya IgM RDT kits received for performance evaluation (Verification/Storage/Unpacking etc).
- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Chikungunya IgM Rapid IVD kit evaluation panel:

Well characterised Chikungunya IVD kit evaluation panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of sera samples should be available from Chikungunya confirmed cases. Further characterised for Chikungunya IgM positivity by using approved reference kits having high sensitivity and specificity.

Chikungunya IgM performance evaluation panel need to be tested again by the reference assays at the time of evaluating a particular index test to confirm the positive and negative status of the samples.

4. Reference assay:

All the samples will be tested by CDC/NIV real-time Chikungunya PCR assay. *Samples which are positive by RT-PCR assay will be further tested by any two of the following IgM ELISA kits:*

i. *ICMR-NIV MAC ELISA kit*
ii. *Inbios* CHIKjj Detect™ IgM ELISA
iii. Anti-Chikungunya virus ELISA (IgM) Test (Euroimmun, Luebeck, Germany)
iv. Any WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved IgM ELISA (as and when available)

Samples positive by PCR and at least two kits will be considered positive. If sufficient RT-PCR positive samples are not available, samples positive by at least 2 ELISA kits (of the kits mentioned above) can be considered as true positive samples.

Samples which are negative by RT-PCR and at least two IgM ELISA kits mentioned above will be considered as Chikungunya negative samples.

5. Sample size and sample panel composition: Sample sizes of positive and negative samples and sample panel composition against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Positive samples: Positive samples should be positive by RT-PCR at least two ELISA kits from the three mentioned above. If sufficient RT-PCR positive samples are not available, samples positive by at least 2 ELISA kits (of the kits mentioned above) can be considered as true positive samples.

Negative samples: Samples which are negative by RT-PCR and at least two IgM ELISA kits mentioned above will be considered as Chikungunya negative samples.

Table 1. Sample sizes and panel composition of positive chikungunya samples for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%	16	20	Strong Positive: 6 Moderate Positive: 7 Weak Positive: 7
95%	77	80	Strong Positive: 24 Moderate Positive: 28 Weak Positive: 28
90%	145	150	Strong Positive: 44 Moderate Positive: 53 Weak Positive: 53
85%	206	210	Strong Positive: 62

			Moderate Positive: 74 Weak Positive: 74
80%	258	260	Strong Positive: 78 Moderate Positive: 91 Weak Positive: 91
The samples need to be classified as strong, moderate and weak positives based on ELISA units of the reference assay. *#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.*			

Table 2. Sample sizes and panel composition of negative chikungunya samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%	16	20	1. Cross-reactivity panel • Dengue IgM positive: 3 @ • Rubella IgM positive: 1 * 2. [a]Acute febrile illness cases: 12 3. [b]Healthy subjects from endemic regions: 4
95%	77	80	1. Cross-reactivity panel • Dengue IgM positive: 13 @ • Rubella IgM positive: 3 * 2. [a]Acute febrile illness cases: 48 3. [b]Healthy subjects from endemic regions: 16
90%	145	150	1. Cross-reactivity panel • Dengue IgM positive: 25 @ • Rubella IgM positive: 5 * 2. [a]Acute febrile illness cases: 90 3. [b]Healthy subjects from endemic regions: 30
85%	206	210	1. Cross-reactivity panel • Dengue IgM positive: 35 @ • Rubella IgM positive: 7 * 2. [a]Acute febrile illness cases: 126 3. [b]Healthy subjects from endemic regions: 42
80%	258	260	1. Cross-reactivity panel • Dengue IgM positive: 43 @ • Rubella IgM positive: 9 * 2. [a]Acute febrile illness cases: 156

			3. [b]Healthy subjects from endemic regions: 52

[a] Acute febrile illness cases negative for above pathogens AND Chikungunya IgM & PCR
[b] Samples from healthy subjects from endemic regions negative for all Chikungunya markers (IgM, RNA)

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

6. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

7. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

8. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

9. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

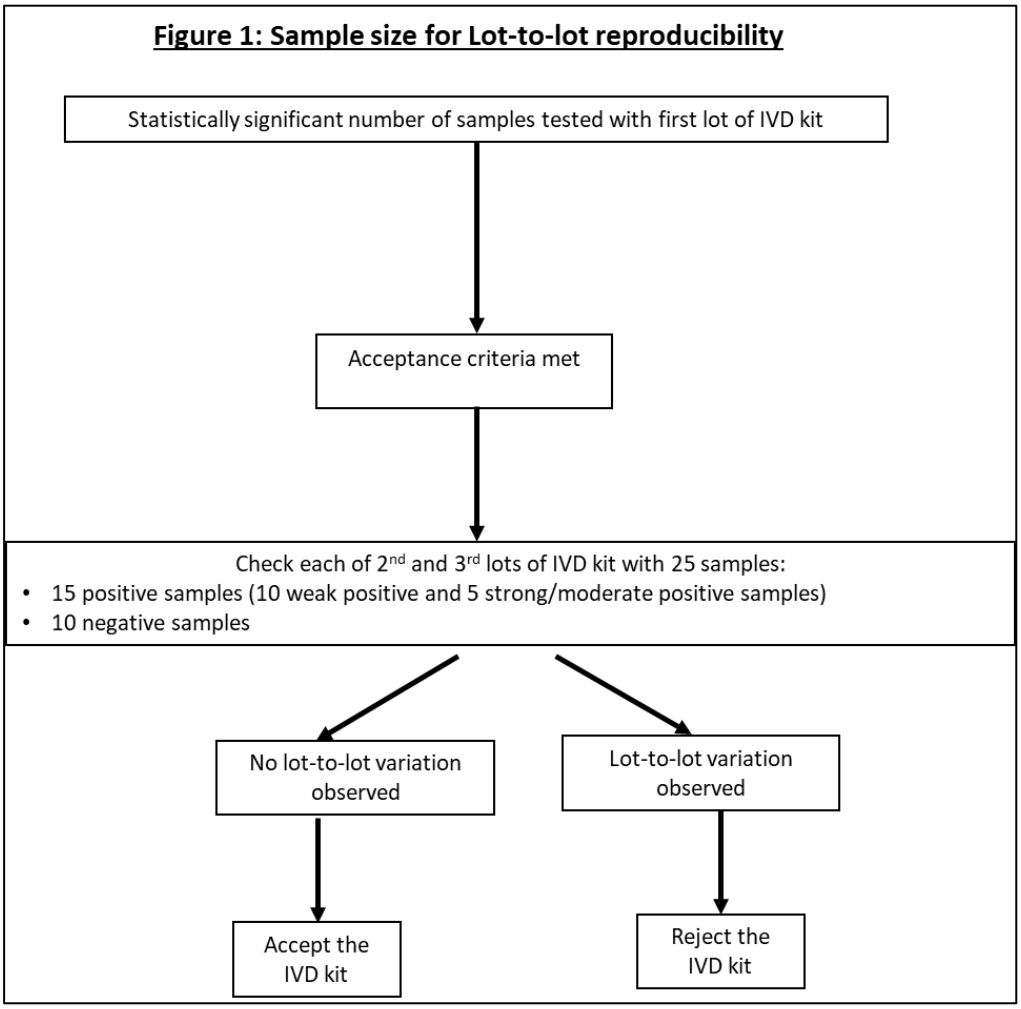
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing and result interpretation should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples

(strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

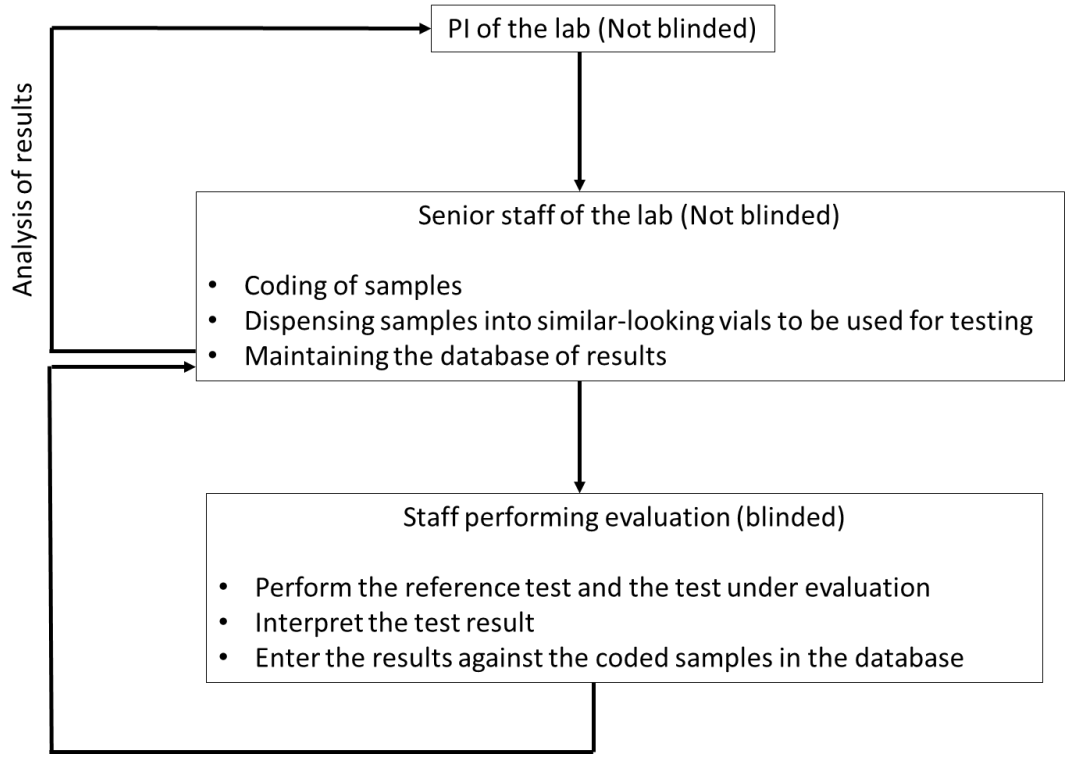
c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

10. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



11. Acceptance criteria:

Sensitivity: ≥80%

Specificity: ≥90%

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥260 positive samples and ≥150 negative samples should be used for evaluation.

12. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Kikuti M, Tauro LB, Moreira PSS, et al. Evaluation of two commercially available Chikungunya virus IgM enzyme-linked immunoassays (ELISA) in a setting of concomitant transmission of Chikungunya, Dengue and Zika viruses. Int J Infect Dis. 2020 Feb;91:38-43.
2. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
3. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR CHIKUNGUNYA IgM RDT KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak/simulated samples)	
	Negative samples (provide details: clinical/spiked, including cross reactivity panel/simulated samples)	

Results:

		Reference assay …….…… (name)		
		Positive	Negative	Total
Name of Chikungunya IgM RDT kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

Conclusions:

- Cross-reactivity:
- Invalid test rate:
- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ..... sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for ……Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……..

Performance evaluation protocol for Chikungunya real-time PCR kits

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Chikungunya PCR kits in the diagnosis of Chikungunya infection using irreversibly de-identified leftover archived/ spiked clinical samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through

A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

- Preparation & characterization of kit evaluation panel

- Handling of Chikungunya PCR kits received for performance evaluation (Verification/Storage/Unpacking etc).
- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Chikungunya RNA evaluation panel:

Well characterised Chikungunya sample panel positive for RNA is a critical requirement for performance evaluation of IVD kits utilizing genome detection. Hence statistically significant number of sera/whole blood samples should be available from Chikungunya PCR confirmed cases.

4. RNA extraction:

RNA extraction should be performed as per reference test and index test Instructions for Use (IFU).

If the manufacturer of the index test recommends a specific RNA extraction kit/system, the same needs to be provided by the manufacturer if the evaluating lab is unable to procure the same/it is not available within the lab's IVD evaluation scope.

5. Real-Time PCR System:

PCR should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Real-time closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.

6. Internal control/Extraction control:

The test under evaluation should have an internal control or extraction control (RNA added before extraction to a sample).

7. Reference assay:

Any WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Chikungunya PCR assay or CDC/NIV protocol for detection of Chikungunya RNA should be used as the reference assay.

All positive samples should be confirmed positive for Chikungunya by reference assay.

All negative samples should be negative for all markers of Chikungunya infection (RNA using reference assay AND IgM using any two of the following kits - ICMR-NIV MAC ELISA kit/Inbios CHIKjj Detect™ IgM ELISA/Anti-Chikungunya virus ELISA (IgM) Test (Euroimmun, Luebeck, Germany)/ any WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved IgM ELISA (as and when available)

8. Sample size and sample panel composition: Sample sizes of positive and negative samples and sample panel composition against different values of sensitivity and

specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\,(1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\,(1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Table 1. Sample sizes and panel composition of positive chikungunya samples for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off]#*	*Sample Panel Composition*
99%	16	20	Strong Positive: 6 Moderate Positive: 7 Weak Positive: 7
95%	77	80	Strong Positive: 24 Moderate Positive: 28 Weak Positive: 28
90%	145	150	Strong Positive: 44 Moderate Positive: 53

			Weak Positive: 53
85%	206	210	Strong Positive: 62 Moderate Positive: 74 Weak Positive: 74
#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.			

Strong positive (Ct value between <25)

Moderate positive (Ct value between 25-30)

Weak positive (Ct value between >30 to 34)

Table 2. Sample sizes and panel composition of negative chikungunya samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%	16	20	1. Cross-reactivity panel • Dengue PCR positive: 4 @ • Rubella PCR positive: 1 * 2. [a]Acute febrile illness cases: 10 3. [b]Healthy subjects from endemic regions: 5
95%	77	80	1. Cross-reactivity panel • Dengue PCR positive: 15 @ • Rubella PCR positive: 5 * 2.[a]Acute febrile illness cases: 40 3.[b]Healthy subjects from endemic regions: 20
90%	145	150	1. Cross-reactivity panel • Dengue PCR positive: 28 @ • Rubella PCR positive: 9 * 2. [a]Acute febrile illness cases: 75 3. [b]Healthy subjects from endemic regions: 38
85%	206	210	1. Cross-reactivity panel • Dengue PCR positive: 39 @ • Rubella PCR positive: 13 * 2. [a]Acute febrile illness cases: 105 3.[b]Healthy subjects from endemic regions: 53
[a] Acute febrile illness cases negative for above pathogens **AND** Chikungunya IgM & PCR [b] Samples from healthy subjects from endemic regions negative for all Chikungunya markers (IgM, RNA).			

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

9. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

10. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

11. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

12. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

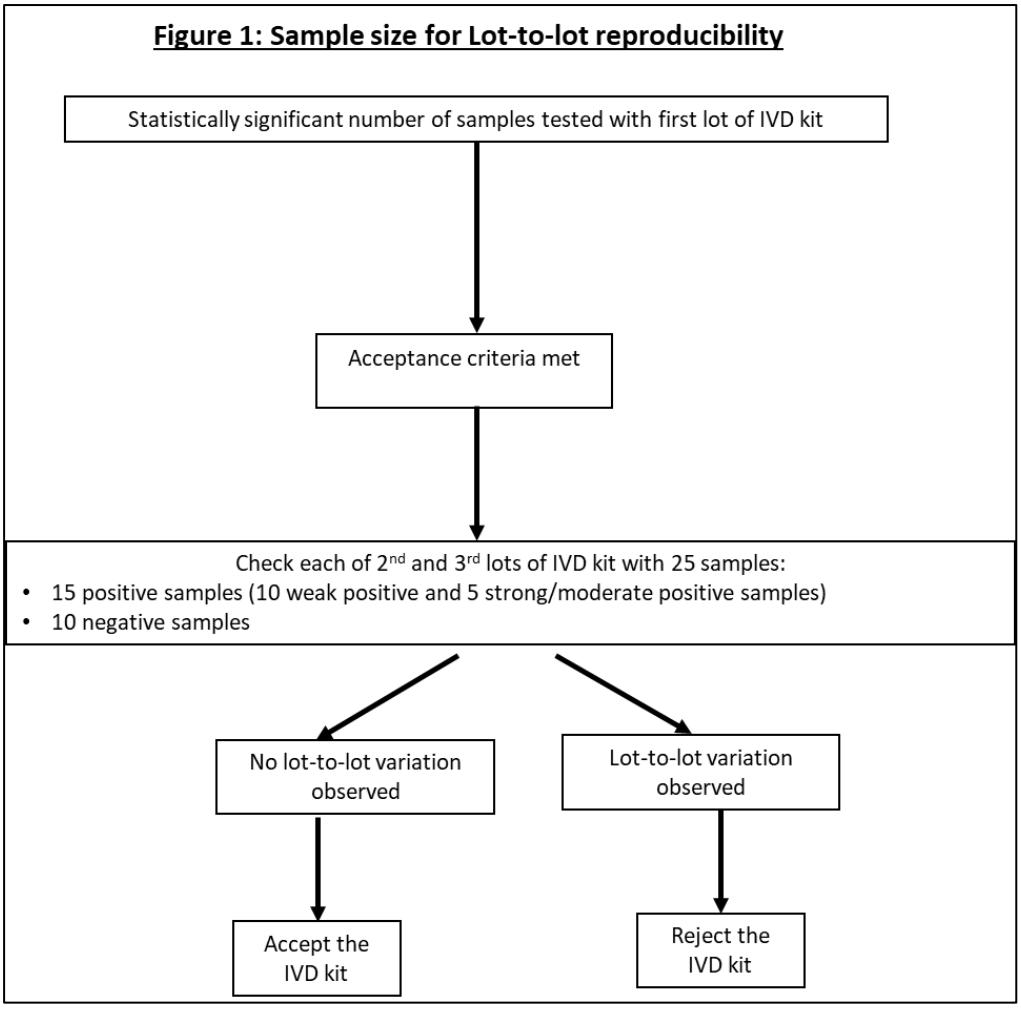
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

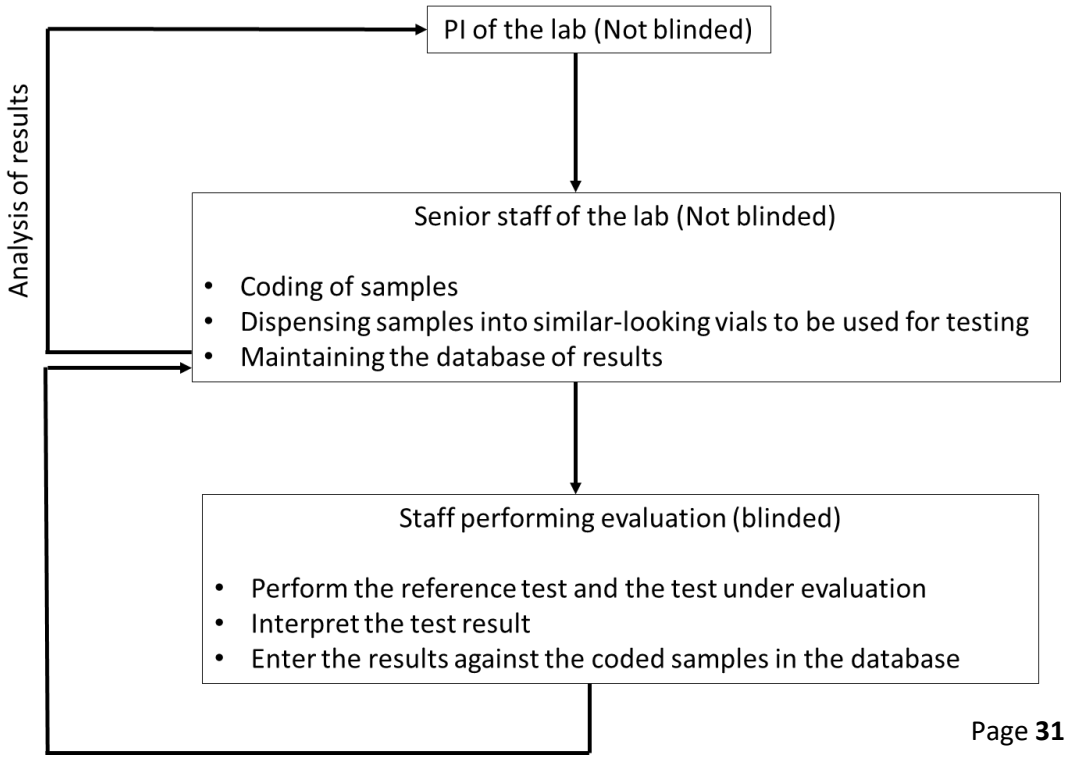
d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report Ct standard deviation across repeats (≤1cycle deviation recommended).

13. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



14. Acceptance criteria:

Sensitivity: ≥95%

Specificity: ≥98%

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥80 positive samples and ≥80 negative samples should be used for evaluation.

15. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Santiago, G.A., Vázquez, J., Courtney, S. et al. Performance of the Trioplex real-time RT-PCR assay for detection of Zika, Dengue, and Chikungunya viruses. Nat Commun 9, 1391 (2018). https://doi.org/10.1038/s41467-018-03772-1
2. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
3. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR CHIKUNGUNYA REAL-TIME PCR KITS

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: clinical/spiked, strong, moderate, weak/simulated samples)	
	Negative samples (provide details: clinical/spiked, including cross reactivity panel/simulated samples)	

Results

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Chikungunya real-time PCR kits	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:
- **Performance: Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ..... sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for Chikungunya…….. Kit (Lot No……) manufactured by …… (supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……..

Performance evaluation protocol for Dengue NS1 RDT kits

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim IVD performance.

II. Purpose:

To evaluate the performance characteristics of Dengue NS1 RDT kits in the diagnosis of Dengue infection using irreversibly de-identified leftover archived/ spiked clinical samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through

A. Accreditation for one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

- Preparation & characterization of kit evaluation panel
- Handling of Dengue NS1 Rapid IVD kits received for performance evaluation (Verification/Storage/Unpacking etc).

- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Dengue RDT IVD kit evaluation panel:

Well characterised Dengue NS1 RDT IVD kit evaluation panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of sera samples should be available from Dengue confirmed cases. Further characterised for Dengue NS1 positivity by using approved reference kits having high sensitivity and specificity.

Dengue NS1 performance evaluation panel need to be tested again by the reference assays at the time of evaluating a particular index test to confirm the positive and negative status of the samples.

4. Reference assay:

WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Dengue NS1 ELISA kit should be used as reference assay.

Serotype status to be assessed using CDC/NIV real-time PCR serotyping protocols.

5. Sample size and sample panel composition: Sample sizes of positive and negative samples and sample panel composition against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\,(1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\,(1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*

- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Positive samples: The panel of positive samples should include samples positive by the reference assay and real-time PCR assay (True positives). Samples should be representative of all 4 serotypes and varying degrees of positivity. The samples should be classified as strong, moderate and weak positives based on ELISA units of the reference assay.

Negative samples: These should include samples negative by the reference NS1 ELISA assay and real-time PCR using CDC/NIV serotyping protocol (True negatives).

Table 1. Sample sizes and panel composition of positive Dengue samples for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%	16	20	Strong Positive: 6 Moderate Positive: 7 Weak Positive: 7
95%	77	80	Strong Positive: 24 Moderate Positive: 28 Weak Positive: 28
90%	145	150	Strong Positive: 44 Moderate Positive: 53 Weak Positive: 53
85%	206	210	Strong Positive: 62 Moderate Positive: 74 Weak Positive: 74
80%	258	260	Strong Positive: 78 Moderate Positive: 91 Weak Positive: 91

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes and panel composition of negative Dengue samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required*	*Sample Panel Composition*

		[Sample size rounded off] #	
99%	16	20	1.<u>PCR/RT-PCR positive samples from other acute febrile illness cases</u> • Chikungunya positive: 4 • Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR): 8 2.<u>Cross-reactivity panel (Samples from other flavivirus disease cases)</u> • Japanese Encephalitis PCR/antigen positive: 1@ • Zika Virus PCR/antigen positive: 1* • West Nile Virus PCR/antigen positive: 1* 3.Samples from healthy subjects from endemic regions negative for all Dengue markers (NS1, IgM, IgG, nucleic acid): 5
95%	77	80	1.<u>PCR/RT-PCR positive samples from other acute febrile illness cases</u> • Chikungunya positive: 15 • Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR): 30 <u>2. Cross-reactivity panel (Samples from other flavivirus disease cases)</u> • Japanese Encephalitis PCR/antigen positive: 5 @ • Zika Virus PCR/antigen positive: 5 * • West Nile Virus PCR/antigen positive: 5 * 3. Samples from healthy subjects from endemic regions negative for all Dengue markers (NS1, IgM, IgG, nucleic acid): 20
90%	145	150	1.<u>PCR/RT-PCR positive samples from other acute febrile illness cases</u> • Chikungunya positive: 28 • Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR): 57 <u>2. Cross-reactivity panel (Samples from other flavivirus disease cases)</u>

			• Japanese Encephalitis PCR/antigen positive: 9@ • Zika Virus PCR/antigen positive: 9 * • West Nile Virus PCR/antigen positive: 9 * 3.Samples from healthy subjects from endemic regions negative for all Dengue markers (NS1, IgM, IgG, nucleic acid): 38
85%	206	210	1.PCR/RT-PCR positive samples from other acute febrile illness cases • Chikungunya positive: 39 • Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR): 79 2. Cross-reactivity panel (Samples from other flavivirus disease cases) • Japanese Encephalitis PCR/antigen positive: 13 @ • Zika Virus PCR/antigen positive: 13 * • West Nile Virus PCR/antigen positive 13 * 3.Samples from healthy subjects from endemic regions negative for all Dengue markers (NS1, IgM, IgG, nucleic acid): 53
80%	258	260	1. PCR/RT-PCR positive samples from other acute febrile illness cases • Chikungunya positive: 49 • Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR): 98 2. Cross-reactivity panel (Samples from other flavivirus disease cases) • Japanese Encephalitis PCR/antigen positive: 16 @ • Zika Virus PCR/antigen positive: 16 * • West Nile Virus PCR/antigen positive: 16 * 3. Samples from healthy subjects from endemic regions negative for all Dengue markers (NS1, IgM, IgG, nucleic acid): 65

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

In the absence of natural samples, spiked samples may be used, as per details provided in the note below.

Commercially available validated standard panels that are accepted by accreditation agencies can also be used if the clinical sample with required analyte level is not available.
Recombinant NS1 antigen of cross reactive flaviviruses (Zika, West Nile and Japanese Encephalitis viruses) expressed in mammalian cells can be obtained commercially and reconstituted in serum samples (100 ng -1 µg/ml) and diluted in the ratio of 1:2 and used accordingly (at least five dilutions for each virus specific NS1).
Before used for evaluation, flavivirus NS1 reconstituted in serum samples needs to be tested by the dengue NS1 reference assay, and dilutions which are negative for dengue should be used for evaluation.
The serum samples used for reconstitution should be negative for Dengue NS1, RNA and IgM antibody.

6. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

7. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

8. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

9. Repeatability and Reproducibility Assessment:

A. <u>Repeatability Assessment</u>

This should be done to assess the repeatability of the detection of target analyte using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

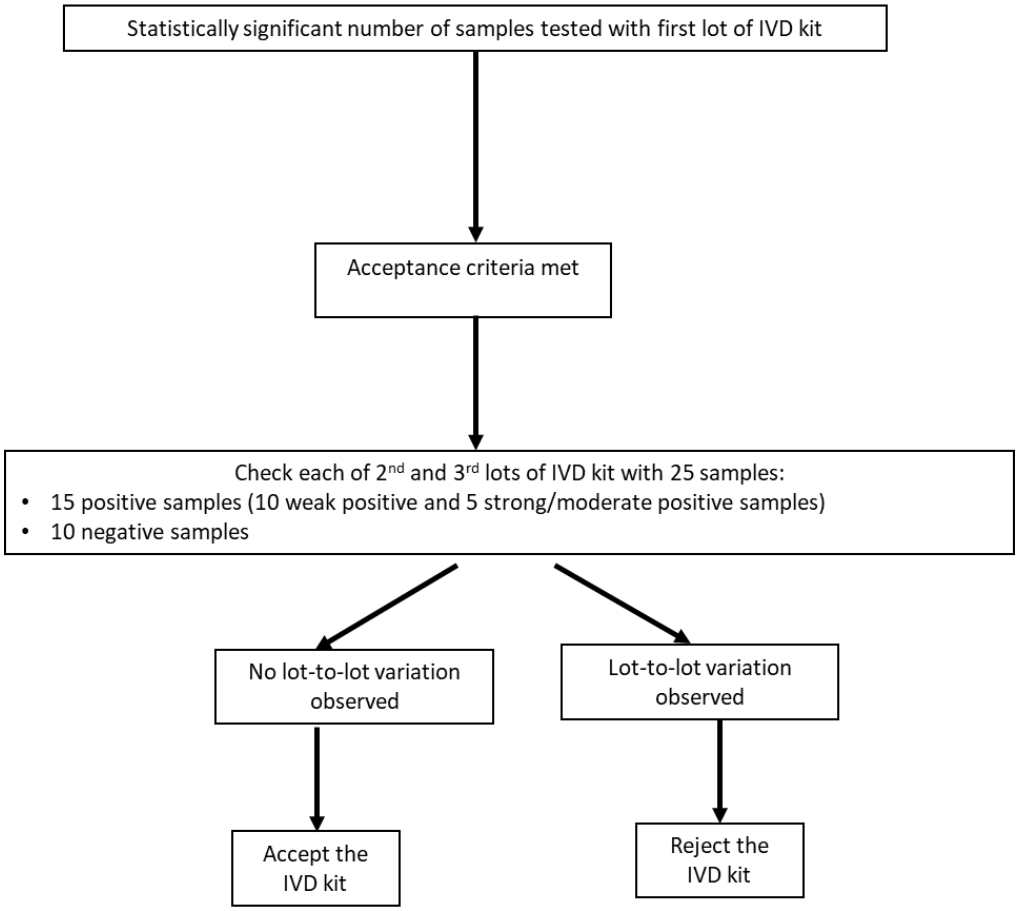
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing and result interpretation should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

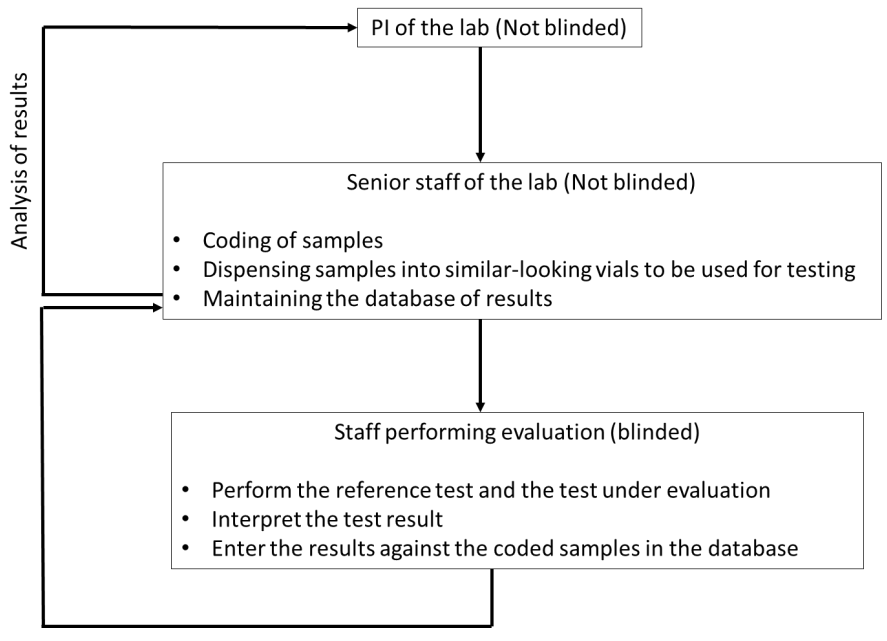
c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

10. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



11. Acceptance criteria:

Sensitivity: ≥80%

Specificity: ≥95%

Cross reactivity with other flavivirus antigens: Minimal

Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥260 positive samples and ≥80 negative samples should be used for evaluation.

12. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, Vazquez S, Cartozian E, Pelegrino JL, Artsob H, Guzman MG, Olliaro P, Zwang J, Guillerm M, Kliks S, Halstead S, Peeling RW, Margolis HS. Evaluation of commercially available diagnostic tests for the detection of Dengue virus NS1 antigen and anti-Dengue virus IgM antibody. PLoSNegl Trop Dis. 2014 Oct 16;8(10):e3171. doi: 10.1371/journal.pntd.0003171.
2. Hermann LL, Thaisomboonsuk B, Poolpanichupatam Y, Jarman RG, Kalayanarooj S, Nisalak A, Yoon IK, Fernandez S. Evaluation of a Dengue NS1 antigen detection assay sensitivity and specificity for the diagnosis of acute Dengue virus infection. PLoSNegl Trop Dis. 2014 Oct 2;8(10):e3193. doi: 10.1371/journal.pntd.0003193.
3. Yow KS, Aik J, Tan EY, Ng LC, Lai YL. Rapid diagnostic tests for the detection of recent Dengue infections: An evaluation of six kits on clinical specimens. PLoS One. 2021 Apr 1;16(4): e0249602. doi: 10.1371/journal.pone.0249602.
4. Mat Jusoh TNA, Shueb RH. Performance Evaluation of Commercial Dengue Diagnostic Tests for Early Detection of Dengue in Clinical Samples. J Trop Med. 2017; 2017: 4687182. doi: 10.1155/2017/4687182. Epub 2017 Dec 12. PMID: 29379526; PMCID: PMC5742879.

5. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
6. Mahajan R, Nair M, Saldanha AM, Harshana A, Pereira AL, Basu N, Goswami RP, Bhattacharya N, Bandyopadhay B, SenGupta M, Day M, Flevaud L, Boelaert M, Burza S. Diagnostic accuracy of commercially available immunochromatographic rapid tests for diagnosis of dengue in India. J Vector Borne Dis. 2021 Apr-Jun;58(2):159-164. doi: 10.4103/0972-9062.321747. PMID: 35074951.
7. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE NS1 RDT KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: clinical/spiked, strong, moderate, weak/simulated samples)	
	Negative samples (provide details: clinical/spiked, including cross reactivity panel/simulated samples)	

Results:

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Dengue NS1 RDT kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:

- **Conclusions:**
  - Sensitivity, specificity
  - Performance: **Satisfactory** / **Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ..... sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for …….. Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……..

Field evaluation protocol for Dengue NS1 RDT kits

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Dengue NS1 RDT kits in the diagnosis of Dengue infection in individuals with unknown disease status.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Laboratory supplies

IV. Ethical approval:

The study will be initiated after approval from the institutional human ethics committee.

V. Procedure:

1. **Study design/type**: Cross-sectional study
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through

   A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel
   - Handling of Dengue NS1 RDT IVD kits received for performance evaluation (Verification/Storage/Unpacking etc).
   - Testing, interpreting, recording of results & reporting
   - Data handling, data safety & confidentiality

3. **Sample size for performance evaluation**:

Sample sizes of positive and negative samples of Dengue against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming

95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. It is further assumed that 30% of the individuals attending the health care facilities for acute febrile illness and suspected for Dengue will be positive for Dengue. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.
Sample size has to be calculated based on both the sensitivity and the specificity. The final sample size will be the maximum of the two. For example, at 95% sensitivity and 95% specificity, the sample size required will be 260 (maximum of 260 and 110).

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x P x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x P x } (1 - \text{IR})}$$

- *n (se) is the minimum number of individuals to be enrolled to obtain the requisite number of positive samples.*
- *n (sp) is the minimum number of individuals to be enrolled to obtain the requisite number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *P is the prevalence of the disease*
- *d is the predetermined marginal error (5%).*
- *IR is the invalid test rate*

Table 1. Sample sizes for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of individuals to be enrolled* [Sample size rounded off] #*
99%	53	60

95%	255	260
90%	484	490
85%	686	690
80%	861	870
* Individuals attending the health care facilities for acute febrile illness and suspected for Dengue meeting the inclusion criteria		

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of individuals to be enrolled[*] [Sample size rounded off] #*
99%	23	30
95%	109	110
90%	207	210
85%	294	300
80%	369	370
* Individuals attending the health care facilities for acute febrile illness and suspected for Dengue meeting the inclusion criteria		

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Recruitment of cases shall be halted once desired number of positive and negative samples are reached.

4. Inclusion criteria:

Individuals with Dengue like illness (An individual with acute febrile illness of 2-7 days with two or more manifestations: Head ache, retro-orbital pain, myalgia, arthralgia, rash, hemorrhagic manifestations)

5. Exclusion criteria:

Individuals with already known positive history for other pathogens

6. Reference assay:

WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Dengue NS1 ELISA kit should be used as reference assay.

Serotype status to be assessed using CDC/NIV real-time PCR serotyping protocols.

7. Study implementation:

The individuals with Dengue like illness will be recruited into the study and five ml of whole blood will be collected in vacutainer tubes and the serum will be separated by centrifugation and used for the study. The serum sample will be subjected to the reference tests and the index test.

8. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

9. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

10. Positive samples:

Samples positive by the reference NS1 ELISA assay and real-time PCR assay will be considered as true positive sample.

11. Negative samples:

Samples negative by the reference *NS1* ELISA assay and real-time PCR using CDC/NIV serotyping protocol will be considered as true negative.

False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

A. Cross reactivity:

Clinical samples or commercially available NS1 antigens from other flaviviruses will be used to test cross reactivity of the index test.

i. Chikungunya PCR positive: 5 samples @
ii. Japanese Encephalitis PCR/antigen positive: 5 samples @
iii. Zika Virus PCR/antigen: 5 samples *
iv. West Nile Virus PCR/antigen: 5 samples *

Cross reactivity panel is arranged in descending order of priority.

The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

In the absence of natural samples, spiked samples may be used, as per details provided in the note below.

Note:

Commercially available validated standard panels that are accepted by accreditation agencies can also be used if the clinical sample with required analyte level is not available. Recombinant NS1 antigen of cross reactive flaviviruses (Zika, West Nile and Japanese Encephalitis viruses) expressed in mammalian cells can be obtained commercially and reconstituted in serum samples (100 ng -1 µg/ml) and diluted in the ratio of 1:2 and used accordingly (at least five dilutions for each virus specific NS1).

Before used for evaluation, flavivirus NS1 reconstituted in serum samples needs to be tested by the dengue NS1 reference assay, and dilutions which are negative for dengue should be used for evaluation.

The serum samples used for reconstitution should be negative for Dengue NS1, RNA and IgM antibody.

12. Statistical analysis:

Sensitivity and specificity will be calculated.

Interim analysis of data shall be conducted on completing evaluation of 25%, 50% and 75% of samples. If, at any point, the performance of the assay is found to be not satisfactory, the assay shall not be evaluated further. Evaluation fee shall be charged accordingly.

13. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

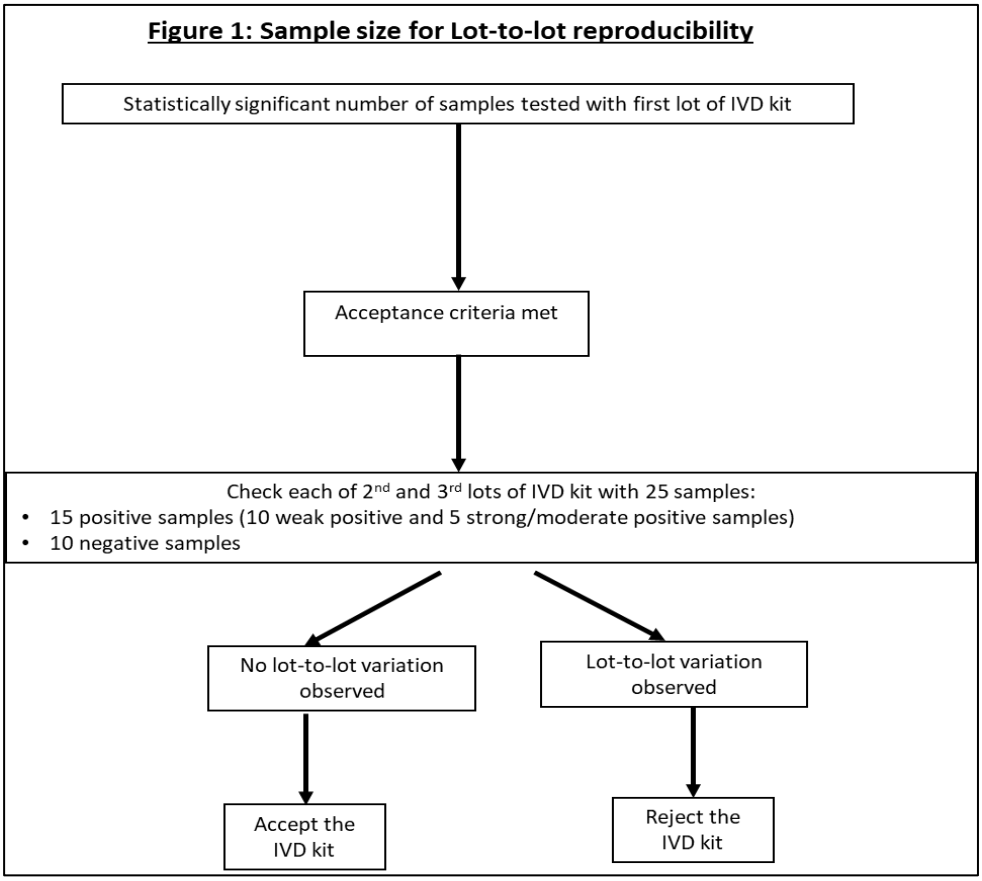
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).

- There should be no lot-to-lot variability (qualitative). Refer Figure 1.



b. Inter-Operator variability: Testing and result interpretation should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

14. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

15. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the results of the reference assay. The PI of the evaluation exercise should remain unblinded, i.e., privy to the results of the reference test. Another senior laboratory staff selected by the PI may remain unblinded for overseeing the activity and maintaining the database of results.. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data.

16. Acceptance Criteria:

Sensitivity: ≥80%
Specificity: ≥95%
Cross-reactivity with other pathogens listed in the negative sample panel: Minimal
Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥870 individuals satisfying the case definition need to be enrolled to obtain the requisite number of positive samples. This sample size is sufficient for required number of negative samples.

Recruitment should be terminated once the desired number of positive cases is enrolled and tested.

17. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, Vazquez S, Cartozian E, Pelegrino JL, Artsob H, Guzman MG, Olliaro P, Zwang J, Guillerm M, Kliks S, Halstead S, Peeling RW, Margolis HS. Evaluation of commercially available diagnostic tests for the detection of Dengue virus NS1 antigen and anti-Dengue virus IgM antibody. PLoSNegl Trop Dis. 2014 Oct 16;8(10):e3171. doi: 10.1371/journal.pntd.0003171.
2. Hermann LL, Thaisomboonsuk B, Poolpanichupatam Y, Jarman RG, Kalayanarooj S, Nisalak A, Yoon IK, Fernandez S. Evaluation of a Dengue NS1 antigen detection assay sensitivity and specificity for the diagnosis of acute Dengue virus infection. PLoSNegl Trop Dis. 2014 Oct 2;8(10):e3193. doi: 10.1371/journal.pntd.0003193.
3. Ganeshkumar P, Murhekar MV, Poornima V, Saravanakumar V, Sukumaran K, Anandaselvasankar A, John D, Mehendale SM. Dengue infection in India: A systematic review and meta-analysis. PLoSNegl Trop Dis. 2018 Jul 16;12(7):e0006618. doi: 10.1371/journal.pntd.0006618.
4. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
5. Mahajan R, Nair M, Saldanha AM, Harshana A, Pereira AL, Basu N, Goswami RP, Bhattacharya N, Bandyopadhay B, SenGupta M, Day M, Flevaud L, Boelaert M, Burza S. Diagnostic accuracy of commercially available immunochromatographic rapid tests for diagnosis of dengue in India. J Vector Borne Dis. 2021 Apr-Jun;58(2):159-164. doi: 10.4103/0972-9062.321747. PMID: 35074951.
6. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE NS1 RDT KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples: Not applicable, may categorize cases as per duration of illness	
	Negative samples (may categorize as per duration of illness, must include cross reactivity panel)	

Results:

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Dengue NS1 RDT kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:
  - Conclusions:
    - Sensitivity, specificity
    - Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in using kits provided by the manufacturer from the batch mentioned above using ….. sample in ……. (field/controlled lab). Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for NS1……Kit (Lot No……) manufactured by ……. (supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In charge …… Seal ……..

Performance evaluation protocol for Dengue NS1 ELISA kits

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Dengue NS1 ELISA kits in the diagnosis of Dengue infection using irreversibly de-identified leftover archived/ spiked clinical samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through

   A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel
   - Handling of Dengue NS1 ELISA kits received for performance evaluation (Verification/Storage/Unpacking etc).

- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Dengue NS1 ELISA IVD kit evaluation panel:

Well characterised Dengue NS1 ELISA IVD kit evaluation panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of sera samples should be available from Dengue confirmed cases. Further characterised for Dengue NS1 positivity by using approved reference kits having high sensitivity and specificity.

Dengue NS1 performance evaluation panel need to be tested again by the reference assays at the time of evaluating a particular index test to confirm the positive and negative status of the samples.

4. Reference assay:

WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Dengue NS1 ELISA kit should be used as reference assay.

Serotype status to be assessed using CDC/NIV real-time PCR serotyping protocols.

5. Sample size and sample panel composition: Sample sizes of positive and negative samples and sample panel composition against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, and an absolute precision of 5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\,(1 - S_e)}{d^2}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\,(1 - S_p)}{d^2}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*

- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*

Positive samples: The panel of positive samples should include samples positive by the reference assay and real-time PCR assay (True positives). Samples should be representative of all 4 serotypes and varying degrees of positivity. The samples should be classified as strong, moderate and weak positives based on ELISA units of the reference assay.

Negative samples: These should include samples negative by the reference NS1 ELISA assay and real-time PCR using CDC/NIV serotyping protocol (True negatives).

Table 1. Sample sizes and panel composition of positive Dengue samples for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%	15	20	Strong Positive: 6 Moderate Positive: 7 Weak Positive: 7
95%	73	80	Strong Positive: 24 Moderate Positive: 28 Weak Positive: 28
90%	138	140	Strong Positive: 42 Moderate Positive: 49 Weak Positive: 49
85%	196	200	Strong Positive: 60 Moderate Positive: 70 Weak Positive: 70
80%	246	250	Strong Positive: 75 Moderate Positive: 87 Weak Positive: 88

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes and panel composition of negative Dengue samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required [Sample size*	*Sample Panel Composition*

		rounded off] #	
99%	15	20	1. Cross-reactivity panel • Chikungunya PCR positive: 4 @ • Japanese Encephalitis PCR/antigen positive: 1 @ • Zika Virus PCR/antigen positive: 1 * • West Nile Virus PCR/antigen positive: 1 * 2. [a]Acute febrile cases negative for Dengue: 8 3. [b]Healthy subjects from endemic regions: 5
95%	73	80	1. Cross-reactivity panel • Chikungunya PCR positive: 15 @ • Japanese Encephalitis PCR/antigen positive: 5 @ • Zika Virus PCR/antigen positive: 5 * • West Nile Virus PCR/antigen positive: 5 * 2.[a]Acute febrile cases negative for Dengue: 30 3.[b]Healthy subjects from endemic regions: 20
90%	138	140	1. Cross-reactivity panel • Chikungunya PCR positive: 26 @ • Japanese Encephalitis PCR/antigen positive: 9 @ • Zika Virus PCR/antigen positive: 9 * • West Nile Virus PCR/antigen positive: 9 * 2.[a]Acute febrile cases negative for Dengue: 52 3.[b]Healthy subjects from endemic regions: 35
85%	196	200	1. Cross-reactivity panel • Chikungunya PCR positive: 37 @ • Japanese Encephalitis PCR/antigen positive: 13 @ • Zika Virus PCR/antigen positive: 13 * • West Nile Virus PCR/antigen positive:13 * 2.[a]Acute febrile cases negative for Dengue: 74 3.[b]Healthy subjects from endemic regions: 50
80%	246	250	1. Cross-reactivity panel • Chikungunya PCR positive: 46 @ • Japanese Encephalitis PCR/antigen positive: 16 @ • Zika Virus PCR/antigen positive: 16 * • West Nile Virus PCR/antigen positive: 16 * 2.[a]Acute febrile cases negative for Dengue: 94 3.[b]Healthy subjects from endemic regions: 62

[a] Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR)
[b] Samples from healthy subjects from endemic regions negative for all Dengue markers (NS1, IgM, IgG, nucleic acid)

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

*In the absence of natural samples, spiked samples may be used, as per details provided in the note below.

Note:

Commercially available validated standard panels that are accepted by accreditation agencies can also be used if the clinical sample with required analyte level is not available. Recombinant NS1 antigen of cross reactive flaviviruses (Zika, West Nile and Japanese Encephalitis viruses) expressed in mammalian cells can be obtained commercially and reconstituted in serum samples (100 ng -1 µg/ml) and diluted in the ratio of 1:2 and used accordingly (at least five dilutions for each virus specific NS1).

Before used for evaluation, flavivirus NS1 reconstituted in serum samples needs to be tested by the dengue NS1 reference assay, and dilutions which are negative for dengue should be used for evaluation.

The serum samples used for reconstitution should be negative for Dengue NS1, RNA and IgM antibody.

6. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

7. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

8. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

9. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

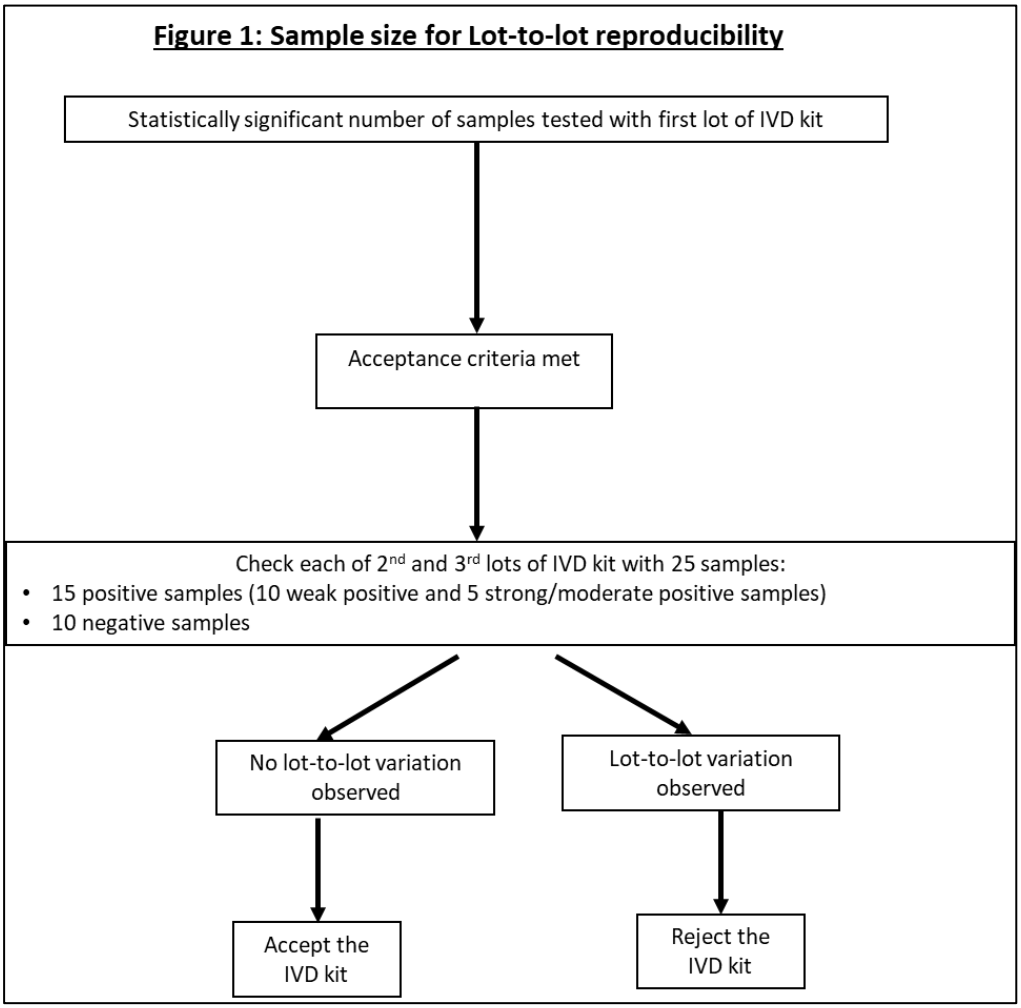
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate

and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

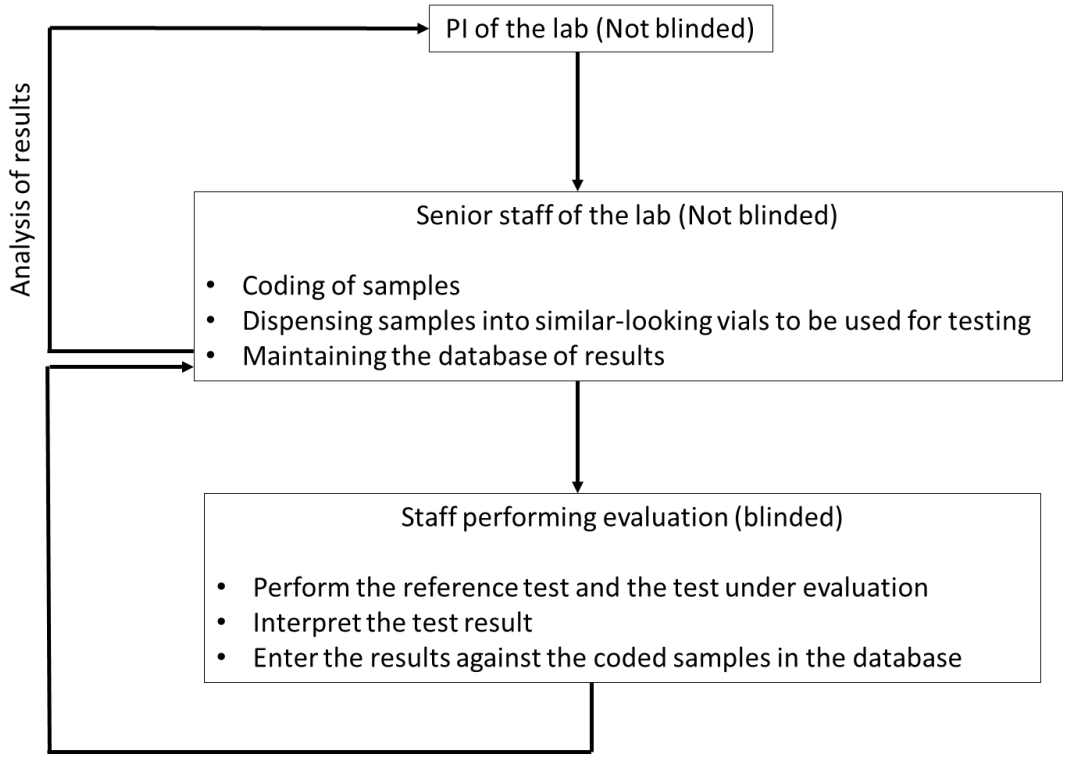
c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

10. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



11. Acceptance Criteria:

Sensitivity: ≥90%

Specificity: ≥95%

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥140 positive samples and ≥80 negative samples should be used for evaluation.

12. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, Vazquez S, Cartozian E, Pelegrino JL, Artsob H, Guzman MG, Olliaro P, Zwang J, Guillerm M, Kliks S, Halstead S, Peeling RW, Margolis HS. Evaluation of commercially available diagnostic tests for the detection of Dengue virus NS1 antigen and anti-Dengue virus IgM antibody. PLoS Negl Trop Dis. 2014 Oct 16;8(10):e3171. doi: 10.1371/journal.pntd.0003171.
2. Hermann LL, Thaisomboonsuk B, Poolpanichupatam Y, Jarman RG, Kalayanarooj S, Nisalak A, Yoon IK, Fernandez S. Evaluation of a Dengue NS1 antigen detection assay sensitivity and specificity for the diagnosis of acute Dengue virus infection. PLoS Negl Trop Dis. 2014 Oct 2;8(10):e3193. doi: 10.1371/journal.pntd.0003193.
3. Central Drugs Standard Control Organization. In-Vitro Diagnostic (IVD) Medical Devices Frequently Asked Questions. 2022. Available at: https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/IVD/FAQs/CDSCO-IVD-FAQ-03-2022-.pdf
4. U.S. Food and Drug Administration. Dengue Virus Serological Reagents - Class II Special Controls Guideline for Industry and Food and Drug Administration Staff. 2014. Available at: https://www.fda.gov/medical-devices/guidance-documents-medical-devices-and-radiation-emitting-products/Dengue-virus-serological-reagents-class-ii-special-controls-guideline-industry-and-food-and-drug

5. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
6. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE NS1 ELISA KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak/simulated samples)	
	Negative samples (provide details: clinical/spiked, including cross reactivity panel/simulated samples)	

Results

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Dengue NS1 ELISA kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:

- **Conclusions:**
  - Sensitivity, specificity

- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ..... sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for ……. Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……

Field evaluation protocol for Dengue NS1 ELISA kits

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Dengue NS1 ELISA kits in the diagnosis of Dengue infection in individuals with unknown disease status.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Laboratory supplies

IV. Ethical approval:

The study will be initiated after approval from the institutional human ethics committee.

V. Procedure:

1. **Study design/type**: Cross-sectional study
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through
   A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel
   - Handling of Dengue NS1 ELISA kits received for performance evaluation (Verification/Storage/Unpacking etc).
   - Testing, interpreting, recording of results & reporting
   - Data handling, data safety & confidentiality

3. Sample size for performance evaluation:

Sample sizes of positive and negative samples of Dengue against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming

95% level of significance, and an absolute precision of 5%. It is further assumed that 30% of the individuals attending the health care facilities for acute febrile illness and suspected for Dengue will be positive for Dengue. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.
Sample size has to be calculated based on both the sensitivity and the specificity. The final sample size will be the maximum of the two. For example, at 95% sensitivity and 95% specificity, the sample size required will be 245 (maximum of 245 and 105).

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x P}}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x P}}$$

- *n (se) is the minimum number of number of individuals to be enrolled to obtain the requisite number of positive samples.*
- *n (sp) is the minimum number of number of individuals to be enrolled to obtain the requisite number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *P is the prevalence of the disease*
- *d is the predetermined marginal error (5%).*

Table 1. Sample sizes for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of individuals to be enrolled* [Sample size rounded off] #*
99%	51	55
95%	243	245
90%	461	465

85%	653	655
80%	820	820
* Individuals attending the health care facilities for acute febrile illness and suspected for Dengue meeting the inclusion criteria		

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of individuals to be enrolled*[*] *[Sample size rounded off] #*
99%	22	25
95%	104	105
90%	198	200
85%	280	280
80%	351	355
* Individuals attending the health care facilities for acute febrile illness and suspected for Dengue meeting the inclusion criteria		

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Recruitment of cases shall be halted once desired number of positive and negative samples are reached.

4. Inclusion criteria:

Individuals with Dengue like illness (A patient with acute febrile illness of 2-7 days with two or more manifestations: Head ache, retro-orbital pain, myalgia, arthralgia, rash, hemorrhagic manifestations)

5. Exclusion criteria:

Individuals with already known positive history for other pathogens

6. Reference assay:

WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Dengue NS1 ELISA kit should be used as reference assay.

Serotype status to be assessed using CDC / NIV real-time PCR serotyping protocols.

7. Study implementation:

The individuals with Dengue like illness will be recruited into the study and five ml of whole blood will be collected in vacutainer tubes and the serum will be separated by centrifugation and used for the study. The serum sample will be subjected to the following reference tests and the index test.

8. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

9. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

10. Positive samples:

Samples positive by the reference NS1 ELISA assay and real-time PCR assay (True positives). will be considered as true positive sample.

11. Negative samples:

Samples negative by the reference *NS1* ELISA assay and real-time PCR using CDC/NIV serotyping protocol will be considered as true negative.

N.B.:

False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

A. Cross reactivity:

Clinical samples or commercially available NS1 antigens from other flaviviruses will be used to test cross reactivity of the index test.

1. Chikungunya PCR positive samples: 5 @
2. Japanese Encephalitis PCR/antigen positive: 5 samples @
3. Zika Virus PCR/antigen: 5 samples *
4. West Nile Virus PCR/antigen: 5 samples *

Cross reactivity panel is arranged in descending order of priority.

The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

In the absence of natural samples, spiked samples may be used, as per details provided in the note below.

Note:

Commercially available validated standard panels that are accepted by accreditation agencies can also be used if the clinical sample with required analyte level is not available. Recombinant NS1 antigen of cross reactive flaviviruses (Zika, West Nile and Japanese Encephalitis viruses) expressed in mammalian cells can be obtained commercially and reconstituted in serum samples (100 ng -1 µg/ml) and diluted in the ratio of 1:2 and used accordingly (at least five dilutions for each virus specific NS1).

Before used for evaluation, NS1 reconstituted in serum samples needs to be tested by the reference assay and dilution which are positive only should be used for evaluation.

The serum samples used for reconstitution should be negative for Dengue NS1, RNA and IgM antibody.

12. Statistical analysis:

Sensitivity and specificity will be calculated.

Interim analysis of data shall be conducted on completing evaluation of 25%, 50% and 75% of samples. If, at any point, the performance of the assay is found to be not satisfactory, the assay shall not be evaluated further. Evaluation fee shall be charged accordingly.

13. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

Reproducibility testing should include the following:

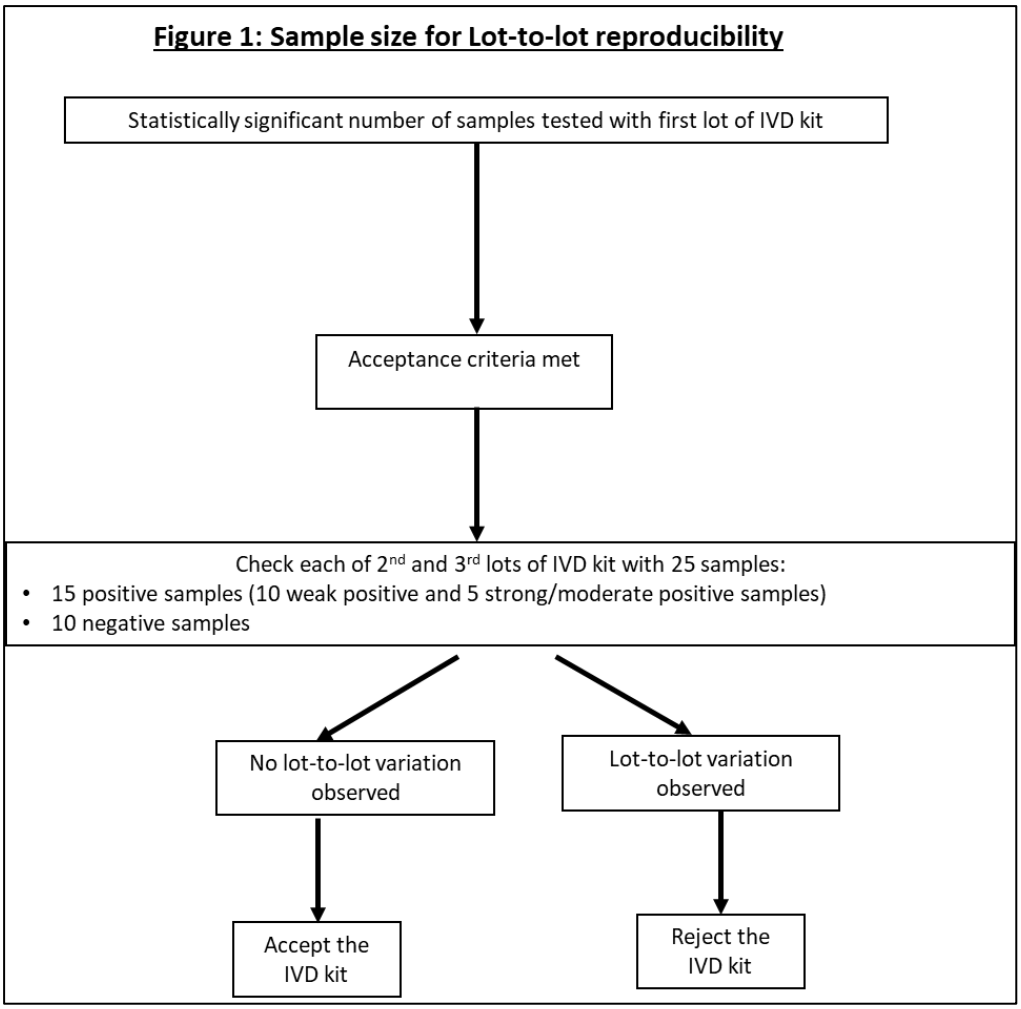
a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).

- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable).

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

14. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

15. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the results of the reference assay. The PI of the evaluation exercise should remain unblinded, i.e., privy to the results of the reference test. Another senior laboratory staff selected by the PI may remain unblinded for overseeing the activity and maintaining the database of results. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data.

16. Acceptance Criteria:

Sensitivity: ≥90%

Specificity: ≥95%

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥465 individuals satisfying the case definition need to be enrolled to obtain the requisite number of positive samples. This sample size is sufficient for required number of negative samples.

Recruitment should be terminated once the desired number of positive cases is enrolled and tested.

17. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, Vazquez S, Cartozian E, Pelegrino JL, Artsob H, Guzman MG, Olliaro P, Zwang J, Guillerm M, Kliks S, Halstead S, Peeling RW, Margolis HS. Evaluation of commercially available diagnostic tests for the detection of Dengue virus NS1 antigen and anti-Dengue virus IgM antibody. PLoSNegl Trop Dis. 2014 Oct 16;8(10):e3171. doi: 10.1371/journal.pntd.0003171.
2. Hermann LL, Thaisomboonsuk B, Poolpanichupatam Y, Jarman RG, Kalayanarooj S, Nisalak A, Yoon IK, Fernandez S. Evaluation of a Dengue NS1 antigen detection assay sensitivity and specificity for the diagnosis of acute Dengue virus infection. PLoSNegl Trop Dis. 2014 Oct 2;8(10):e3193. doi: 10.1371/journal.pntd.0003193.
3. Ganeshkumar P, Murhekar MV, Poornima V, Saravanakumar V, Sukumaran K, Anandaselvasankar A, John D, Mehendale SM. Dengue infection in India: A systematic review and meta-analysis. PLoSNegl Trop Dis. 2018 Jul 16;12(7):e0006618. doi: 10.1371/journal.pntd.0006618.
4. Castro-Jorge LA, Machado PR, Fávero CA, Borges MC, Passos LM, de Oliveira RM, Fonseca BA. Clinical evaluation of the NS1 antigen-capture ELISA for early diagnosis of Dengue virus infection in Brazil. J Med Virol. 2010 Aug;82(8):1400-5. doi: 10.1002/jmv.21814.
5. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
6. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE NS1 ELISA KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples: Not applicable, may categorize cases as per duration of illness	
	Negative samples (may categorize as per duration of illness, must include cross reactivity panel)	

Results

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Dengue NS1 ELISA kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:
  - Conclusions:
    - Sensitivity, specificity
    - Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in using kits provided by the manufacturer from the batch mentioned above using ….. sample in controlled lab setting. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for NS1……Kit (Lot No……) manufactured by …… (supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In charge …… Seal ……..

Performance evaluation protocol for Dengue IgM RDT kits

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Dengue IgM RDT kits in the diagnosis of Dengue infection using irreversibly de-identified leftover archived/ spiked clinical samples..

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through
   A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

   B. Staff training**:** All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel
   - Handling of Dengue IgM Rapid IVD kits received for performance evaluation (Verification/Storage/Unpacking etc).

- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Dengue IgM Rapid IVD kit evaluation panel:

Well characterised Dengue IVD kit evaluation panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of sera samples should be available from Dengue confirmed cases. Further characterised for Dengue IgM positivity by using approved reference kits having high sensitivity and specificity.

Dengue IgM performance evaluation panel need to be tested again by the reference assays at the time of evaluating a particular index test to confirm the positive and negative status of the samples.

4. Reference assay:

WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Dengue IgM ELISA kit should be used as reference assay.

NS1 antigen status to be assessed using WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved NS1 ELISA kit.

Serotype status to be assessed using a combination of CDC/NIV real-time PCR serotyping protocols.

At least 50% of the samples should be positive by real-time PCR or NS1 antigen and IgM ELISA**.**

Primary and Secondary status to be assessed by Panbio Dengue IgG capture ELISA kit.

5. Sample size and sample panel composition: Sample sizes of positive and negative samples of Dengue against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\left(1 - S_p\right)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Positive samples: The panel of positive samples should include samples positive by the reference assay, with 50% samples positive for Dengue NS1/RT-PCR assay (True positives). Samples should be representative of all 4 serotypes and varying degrees of positivity. The samples should be classified as strong, moderate and weak positives based on ELISA units of the reference assay.

Negative samples: These should include samples negative by the reference assay, NS1 ELISA assay and/or real-time PCR using CDC/NIV serotyping protocol (True negatives).

Table 1. Sample sizes and panel composition of positive Dengue samples for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%	16	20	Strong Positive: 6 Moderate Positive: 7 Weak Positive: 7
95%	77	80	Strong Positive: 24 Moderate Positive: 28 Weak Positive: 28
90%	145	150	Strong Positive: 44 Moderate Positive: 53 Weak Positive: 53
85%	206	210	Strong Positive: 62 Moderate Positive: 74 Weak Positive: 74
80%	258	260	Strong Positive: 78

			Moderate Positive: 91 Weak Positive: 91

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes and panel composition of negative Dengue samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required [Sample size rounded off]#*	*Sample Panel Composition*
99%[#]	16	20	1.Cross-reactivity panel • Chikungunya IgM positive: 4 @ • Japanese Encephalitis IgM positive: 1 @ • Zika Virus IgM positive: 1 * • West Nile Virus IgM positive: 1 * 2.[a]Acute febrile cases: 5 3.Rheumatoid Arthritis/other autoimmune disease cases: 4 4.[b]Healthy subjects from endemic regions: 4
95%	77	80	1.Cross-reactivity panel • Chikungunya IgM positive: 16 @ • Japanese Encephalitis IgM positive: 3 @ • Zika Virus IgM positive: 3 * • West Nile Virus IgM positive: 3 * 2.[a]Acute febrile cases: 23 3.Rheumatoid Arthritis/other autoimmune disease cases: 16 4.[b]Healthy subjects from endemic regions: 16
90%	145	150	1.Cross-reactivity panel • Chikungunya IgM positive: 30 @ • Japanese Encephalitis IgM positive: 5 @ • Zika Virus IgM positive: 5 * • West Nile Virus IgM positive: 5 * 2.[a]Acute febrile cases: 45 3.Rheumatoid Arthritis/other autoimmune disease cases: 30 4.[b]Healthy subjects from endemic regions: 30

85%	206	210	1.Cross-reactivity panel • Chikungunya IgM positive: 42 @ • Japanese Encephalitis IgM positive: 7 @ • Zika Virus IgM positive: 7 * • West Nile Virus IgM positive: 7 * 2.[a]Acute febrile cases: 63 3.Rheumatoid Arthritis/other autoimmune disease cases: 42 4.[b]Healthy subjects from endemic regions: 42
80%	258	260	1.Cross-reactivity panel • Chikungunya IgM positive: 52 @ • Japanese Encephalitis IgM positive: 9 @ • Zika Virus IgM positive: 9 * • West Nile Virus IgM positive: 9 * 2.[a]Acute febrile cases: 77 3.Rheumatoid Arthritis/other autoimmune disease cases: 52 4.[b]Healthy subjects from endemic regions: 52

[a] Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR)
[b] Samples from healthy subjects from endemic regions negative for all Dengue markers (NS1, IgM, IgG, RNA)
#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

Note: Commercially available validated standard panels that are accepted by accreditation agencies can also be used if the clinical sample with required analyte level is not available. If IgM positive samples for cross reactive flaviviruses are not available, commercially available IgM sera panel for different viruses can be procured and used to test cross reactivity.

6. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

7. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

8. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

9. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte(s) using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

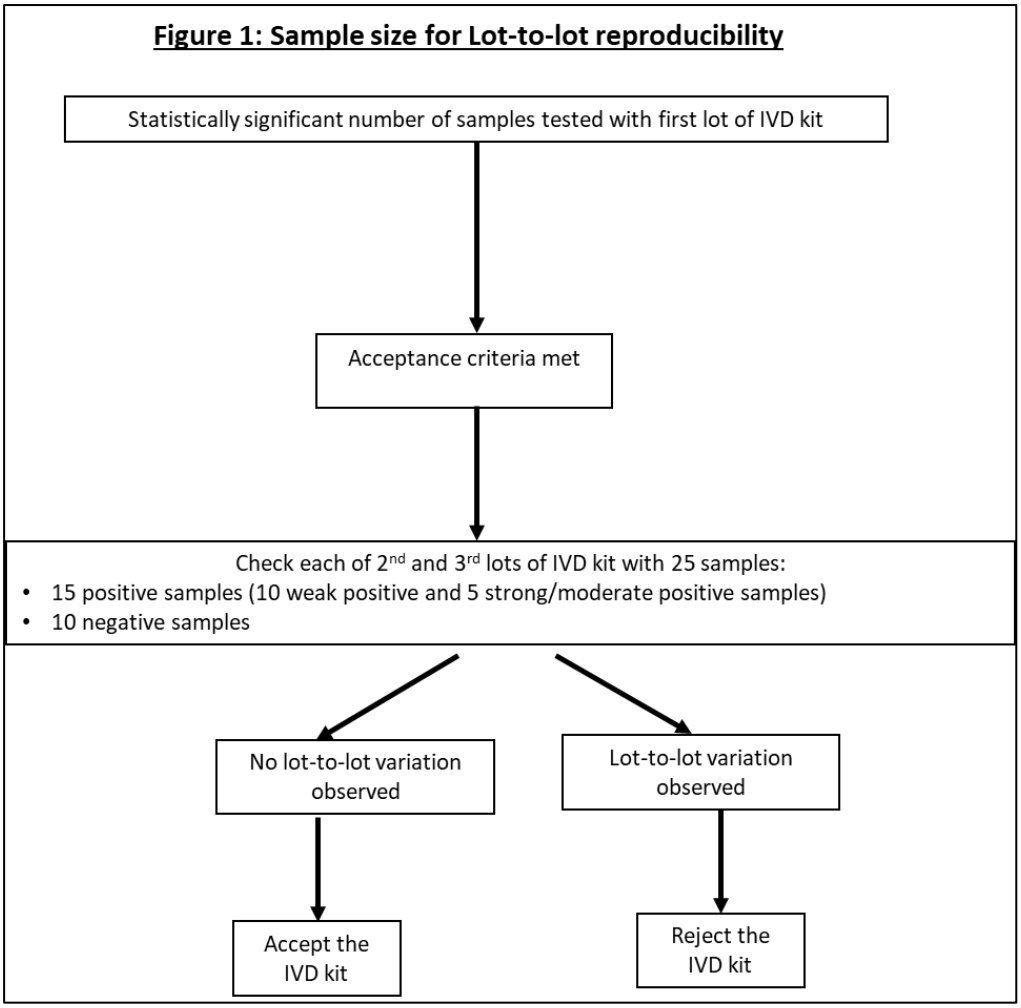
B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.



b. Inter-Operator variability: Testing and result interpretation should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

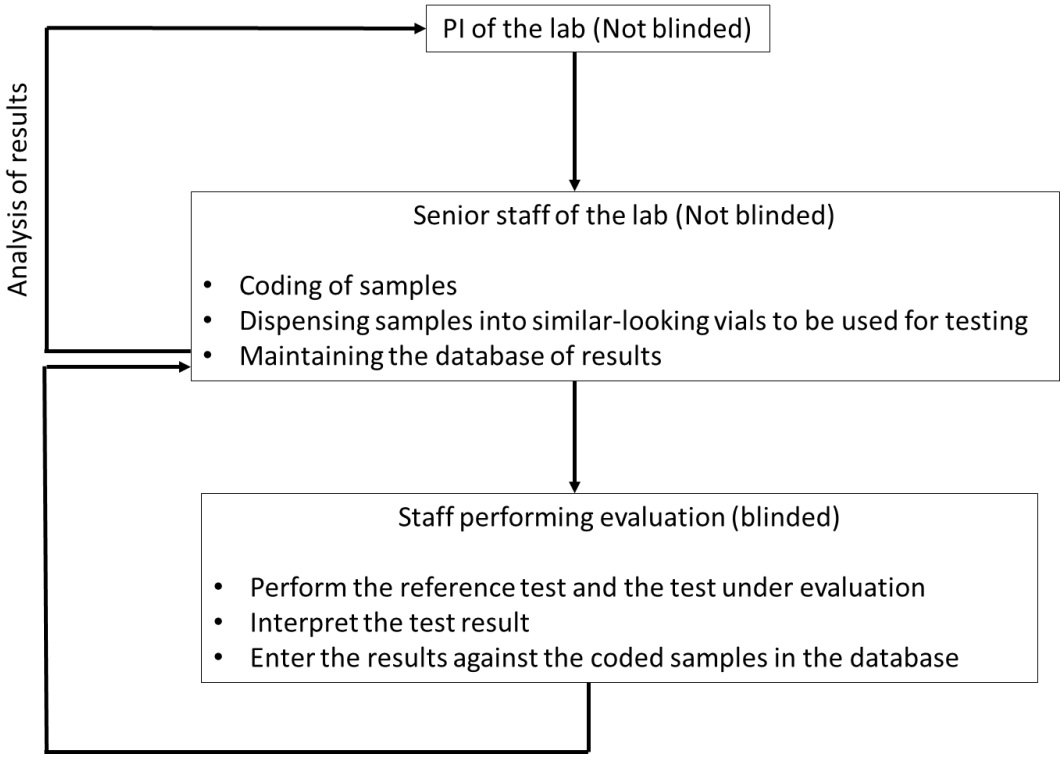
c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

10. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



11. Acceptance Criteria:

Sensitivity: ≥80%

Specificity: ≥90%

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥260 positive samples and ≥150 negative samples should be used for evaluation.

12. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, Pelegrino JL, Vázquez S, Artsob H, Drebot M, Gubler DJ, Halstead SB, Guzmán MG, Margolis HS, Nathanson CM, Rizzo Lic NR, Bessoff KE, Kliks S, Peeling RW. Evaluation of commercially available anti-Dengue virus immunoglobulin M tests. Emerg Infect Dis. 2009 Mar;15(3):436-40. doi: 10.3201/eid1503.080923.
2. Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, Vazquez S, Cartozian E, Pelegrino JL, Artsob H, Guzman MG, Olliaro P, Zwang J, Guillerm M, Kliks S, Halstead S, Peeling RW, Margolis HS. Evaluation of commercially available diagnostic tests for the detection of Dengue virus NS1 antigen and anti-Dengue virus IgM antibody. PLoS Negl Trop Dis. 2014 Oct 16;8(10):e3171. doi: 10.1371/journal.pntd.0003171.
3. WHO, Evaluation of commercially available anti-Dengue virus immunoglobulin M tests. (Diagnostics evaluation series, 3). ISBN 978 92 4 159775 3.
4. Central Drugs Standard Control Organization. Guidance on Performance Evaluation of In-vitro Diagnostic Medical Devices. 2018. Available at: https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/medical device/guidanceperformanceivd.pdf
5. Central Drugs Standard Control Organization. In-Vitro Diagnostic (IVD) Medical Devices Frequently Asked Questions. 2022. Available at: https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/IVD/FAQs/CDSCO-IVD-FAQ-03-2022-.pdf
6. U.S. Food and Drug Administration. Dengue Virus Serological Reagents - Class II Special Controls Guideline for Industry and Food and Drug Administration Staff. 2014. Available at: https://www.fda.gov/medical-devices/guidance-documents-medical-devices-and-radiation-emitting-products/Dengue-virus-serological-reagents-class-ii-special-controls-guideline-industry-and-food-and-drug
7. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1

8. Yow KS, Aik J, Tan EY, Ng LC, Lai YL. Rapid diagnostic tests for the detection of recent Dengue infections: An evaluation of six kits on clinical specimens. PLoS One. 2021 Apr 1;16(4):e0249602. doi: 10.1371/journal.pone.0249602.

9. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

***The validation protocols need to be revisited after introduction of Dengue vaccines and the acceptance criteria needs revisiting every year so as to enable the availability of best diagnostic kits.**

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE IgM RDT KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak/simulated samples)	
	Negative samples (provide details: clinical/spiked, including cross reactivity panel/simulated samples)	

Results:

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Dengue IgM RDT kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:
- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ….. sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for ……Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal …….

Performance evaluation protocol for Dengue IgM ELISA kits

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Dengue IgM ELISA kits in the diagnosis of Dengue infection using irreversibly de-identified leftover archived/ spiked clinical samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approval:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through
   A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel
   - Handling of Dengue IgM ELISA IVD kits received for performance evaluation (Verification/Storage/Unpacking etc).

- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Dengue IgM ELISA IVD kit evaluation panel:

Well characterised Dengue IVD kit evaluation panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of sera samples should be available from Dengue confirmed cases. Further characterised for Dengue IgM positivity by using approved reference kits having high sensitivity and specificity.

Dengue IgM performance evaluation panel need to be tested again by the reference assays at the time of evaluating a particular index test to confirm the positive and negative status of the samples.

4. Reference assay:

WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Dengue IgM ELISA kit should be used as reference assay.

NS1 antigen status to be assessed using WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved NS1 ELISA kit.

Serotype status to be assessed using a combination of CDC/NIV real-time PCR serotyping protocols.

At least 50% of the samples should be positive by real-time PCR or NS1 antigen and IgM ELISA**.**

Primary and Secondary status to be assessed by Panbio Dengue IgG capture ELISA kit**.**

5. Sample size and sample panel composition: Sample sizes of positive and negative samples and sample panel composition against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, and an absolute precision of 5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\,(1 - S_e)}{d^2}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\,(1 - S_p)}{d^2}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*

Positive samples: The panel of positive samples should include samples positive by the reference assay, with 50% samples positive for Dengue NS1/ RT-PCR assay (True positives). Samples should be representative of primary/secondary Dengue and all 4 Dengue virus serotypes, with varying degrees of positivity. The samples should be classified as strong, moderate and weak positives based on ELISA units of the reference assay.

Negative samples: These should include samples negative by the reference assay, NS1 ELISA and/or real-time PCR using CDC and/or NIV serotyping protocols. (True negatives).

Table 1. Sample sizes and panel composition of positive Dengue samples for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off]#*	*Sample Panel Composition*
99%	15	20	Strong Positive: 6 Moderate Positive: 7 Weak Positive: 7
95%	73	80	Strong Positive: 24 Moderate Positive: 28 Weak Positive: 28
90%	138	140	Strong Positive: 42 Moderate Positive: 49 Weak Positive: 49
85%	196	200	Strong Positive: 60 Moderate Positive: 70 Weak Positive: 70
80%	246	250	Strong Positive: 75 Moderate Positive: 87 Weak Positive: 88

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes and panel composition of negative Dengue samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required [Sample size rounded off]#*	*Sample Panel Composition*
99%[#]	15	20	1.Cross-reactivity panel • Chikungunya IgM positive: 3 @ • Japanese Encephalitis IgM positive: 1 @ • West Nile Virus IgM positive: 1 * • Zika Virus IgM positive: 1 * 2.Rheumatoid Arthritis/other autoimmune disease cases: 4 3.[a]Acute febrile cases: 6 4.[b]Healthy subjects from endemic regions: 4
95%	73	80	1.Cross-reactivity panel • Chikungunya IgM positive: 10 @ • Japanese Encephalitis IgM positive: 5 @ • West Nile Virus IgM positive: 5 * • Zika Virus IgM positive: 5 * 2.Rheumatoid Arthritis/other autoimmune disease cases: 15 3.[a]Acute febrile cases: 25 4.[b]Healthy subjects from endemic regions: 15
90%	138	140	1.Cross-reactivity panel • Chikungunya IgM positive: 18 @ • Japanese Encephalitis IgM positive: 9 @ • West Nile Virus IgM positive: 9 * • Zika Virus IgM positive: 9 * 2.Rheumatoid Arthritis/other autoimmune disease cases: 26 3.[a]Acute febrile cases: 43 4.[b]Healthy subjects from endemic regions: 26
85%	196	200	1.Cross-reactivity panel • Chikungunya IgM positive: 25 @ • Japanese Encephalitis IgM positive: 12 @ • West Nile Virus IgM positive: 12 * • Zika Virus IgM positive: 12 * 2.Rheumatoid Arthritis/other autoimmune disease cases: 38 3.[a]Acute febrile cases: 63 4.[b]Healthy subjects from endemic regions: 38

80%	246	250	1.Cross-reactivity panel • Chikungunya IgM positive: 31@ • Japanese Encephalitis IgM positive: 16 @ • West Nile Virus IgM positive: 16 * • Zika Virus IgM positive: 16 * 2.Rheumatoid Arthritis/other autoimmune disease cases: 47 3.[a]Acute febrile cases: 77 4.[b]Healthy subjects from endemic regions: 47

[a] Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR)
[b] Samples from healthy subjects from endemic regions negative for all Dengue markers (NS1, IgM, IgG, RNA)

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

Note: Commercially available validated standard panels that are accepted by accreditation agencies can also be used if the clinical sample with required analyte level is not available. If IgM positive samples for cross reactive flaviviruses are not available, commercially available IgM sera panel for different viruses can be procured and used to test cross reactivity.

6. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

7. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

8. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

9. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

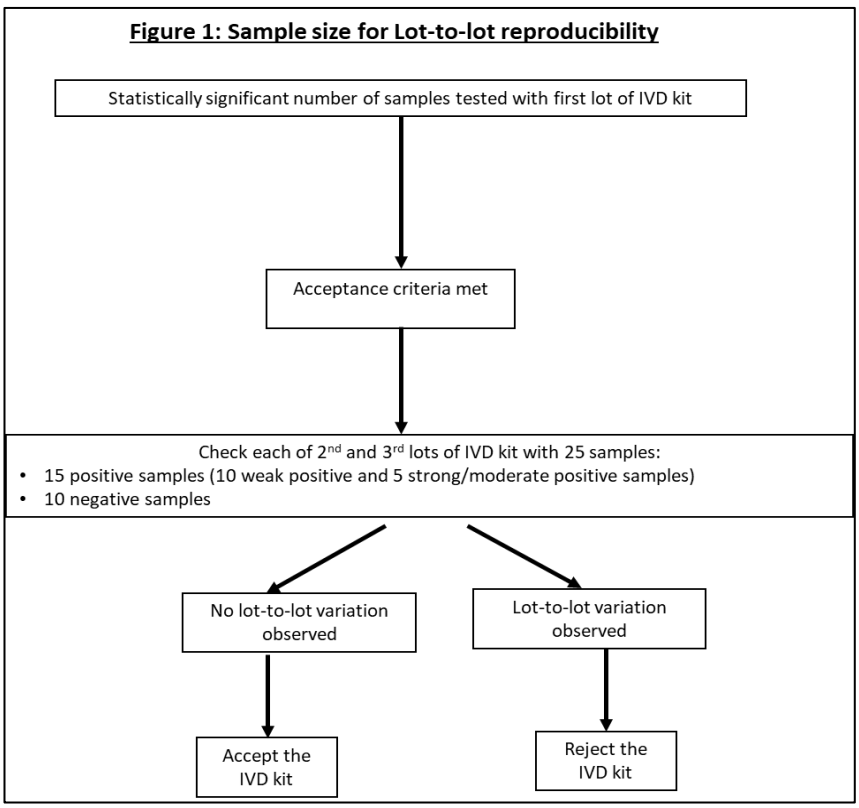
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

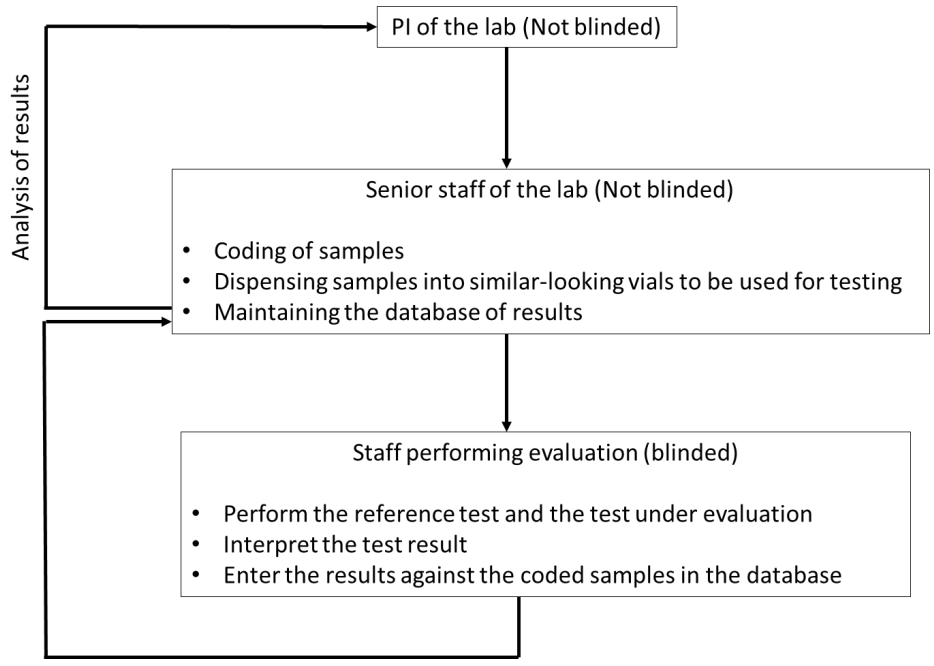
c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

10. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



11. Acceptance criteria

Sensitivity: ≥90%

Specificity: ≥95%

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥140 positive samples and ≥80 negative samples should be used for evaluation.

12. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, Pelegrino JL, Vázquez S, Artsob H, Drebot M, Gubler DJ, Halstead SB, Guzmán MG, Margolis HS, Nathanson CM, Rizzo Lic NR, Bessoff KE, Kliks S, Peeling RW. Evaluation of commercially available anti-Dengue virus immunoglobulin M tests. Emerg Infect Dis. 2009 Mar;15(3):436-40. doi: 10.3201/eid1503.080923.
2. Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, Vazquez S, Cartozian E, Pelegrino JL, Artsob H, Guzman MG, Olliaro P, Zwang J, Guillerm M, Kliks S, Halstead S, Peeling RW, Margolis HS. Evaluation of commercially available diagnostic tests for the detection of Dengue virus NS1 antigen and anti-Dengue virus IgM antibody. PLoSNegl Trop Dis. 2014 Oct 16;8(10):e3171. doi: 10.1371/journal.pntd.0003171.
3. WHO, Evaluation of commercially available anti-Dengue virus immunoglobulin M tests. (Diagnostics evaluation series, 3). ISBN 978 92 4 159775 3.
4. Central Drugs Standard Control Organization. Guidance on Performance Evaluation of In-vitro Diagnostic Medical Devices. 2018. Available at: https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/medical device/guidanceperformanceivd.pdf
5. Central Drugs Standard Control Organization. In-Vitro Diagnostic (IVD) Medical Devices Frequently Asked Questions. 2022. Available at: https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/IVD/FAQs/CDSCO-IVD-FAQ-03-2022-.pdf

6. U.S. Food and Drug Administration. Dengue Virus Serological Reagents - Class II Special Controls Guideline for Industry and Food and Drug Administration Staff. 2014. Available at: https://www.fda.gov/medical-devices/guidance-documents-medical-devices-and-radiation-emitting-products/Dengue-virus-serological-reagents-class-ii-special-controls-guideline-industry-and-food-and-drug
7. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
8. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

***The validation protocols need to be revisited after introduction of Dengue vaccines and the acceptance criteria needs revisiting every year so as to enable the availability of best diagnostic kits.**

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE IgM ELISA KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak/simulated samples)	
	Negative samples (provide details: clinical/spiked, including cross reactivity panel/simulated samples)	

Results:

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Dengue IgM ELISA kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:
- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ….. sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for ……Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal……

Performance evaluation protocol for Dengue NS1/IgM combo RDT kits

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Dengue NS1/IgM combo RDT kits in the diagnosis of Dengue infection.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through
   A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel
   - Handling of Dengue NS1/IgM combo IVD kits received for performance evaluation (Verification/Storage/Unpacking etc).

- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Dengue RDT IVD kit evaluation panel:

Well characterised Dengue RDT IVD kit evaluation panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of sera samples should be available from Dengue confirmed cases. Further characterised for Dengue NS1 and IgM positivity by using approved reference kits having high sensitivity and specificity.

Dengue NS1/IgM performance evaluation panel need to be tested again by the reference assays at the time of evaluating a particular index test to confirm the positive and negative status of the samples.

4. Reference assay:

Anti-DENV IgM detection ELISA WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved kit

AND/OR

DENV NS1 ELISA WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved kit

Serotype status to be assessed using a combination of CDC and/or NIV real-time PCR serotyping protocols.

All positive samples need confirmation reference NS1/IgM ELISA assay and real-time PCR assay.

5. Sample size and sample panel composition: Sample sizes of positive and negative samples of Dengue against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\,(1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\,(1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Positive samples: Samples which are positive for IgM or NS1 or both by the reference assays will be considered as true positive samples. There should be representation of samples positive for all four serotypes.

Negative samples: These should include samples negative by all the reference assays and real-time PCR using CDC and/or NIV serotyping protocol (True negatives).

Table 1. Sample sizes and panel composition of positive Dengue samples for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off for balanced allocation] #*	*Sample Panel Composition*
99%[#]	16	28	[*]NS1 positive and IgM negative: 8 [*]NS1 and IgM positive: 12 [*]NS1 negative and IgM positive: 8
95%	77	84	[*]NS1 positive and IgM negative: 24 [*]NS1 and IgM positive: 36 [*]NS1 negative and IgM positive: 24
90%	145	160	[*]NS1 positive and IgM negative: 44 [*]NS1 and IgM positive: 72 [*]NS1 negative and IgM positive: 44
85%	206	220	[*]NS1 positive and IgM negative: 60 [*]NS1 and IgM positive: 100 [*]NS1 negative and IgM positive: 60
80%	258	260	[*]NS1 positive and IgM negative: 72 [*]NS1 and IgM positive: 116 [*]NS1 negative and IgM positive: 72

*** all 4 serotypes shall be represented**

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Note:
In the absence of natural samples, spiked samples may be used as per details provided below:

Commercially available validated standard panels that are accepted by accreditation agencies can also be used if the clinical sample with required analyte level is not available. Recombinant NS1 antigen of cross reactive flaviviruses (Zika, West Nile and Japanese Encephalitis viruses) expressed in mammalian cells can be obtained commercially and reconstituted in serum samples (100 ng -1 µg/ml) and diluted in the ratio of 1:2 and used accordingly (at least five dilutions for each virus specific NS1).
Before used for evaluation, flavivirus NS1 reconstituted in serum samples needs to be tested by the dengue NS1 reference assay, and dilutions which are negative for dengue should be used for evaluation.
The serum samples used for reconstitution should be negative for Dengue NS1, RNA and IgM antibody.

Table 2. Sample sizes and panel composition of negative Dengue samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required [Sample size rounded off for balanced allocation]#*	*Sample Panel Composition*
99%[#]	16	30	1.Cross-reactivity panel For antibody: • Chikungunya IgM positive: 1 @ • Japanese Encephalitis IgM positive: 1 @ • West Nile Virus IgM positive: 1 * • Zika Virus IgM positive: 1 * For antigen: • Chikungunya PCR positive: 1 @ • Japanese Encephalitis NS1/PCR positive: 1 @ • West Nile Virus NS1/PCR positive: 1 * • Zika Virus NS1/PCR positive: 1 * 2.Rheumatoid Arthritis/other autoimmune disease cases: 5 3.[a]Acute febrile cases: 12 4.[b]Healthy subjects from endemic regions: 5
95%	77	90	1.Cross-reactivity panel For antibody: • Chikungunya IgM positive: 3 @

			• Japanese Encephalitis IgM positive: 3 @ • West Nile Virus IgM positive: 3 * • Zika Virus IgM positive: 3 * <u>For antigen:</u> • Chikungunya PCR positive: 3 @ • Japanese Encephalitis NS1/PCR positive: 3 @ • West Nile Virus NS1/PCR positive: 3 * • Zika Virus NS1/PCR positive: 3 * 2.Rheumatoid Arthritis/other autoimmune disease cases: 15 3.[a]Acute febrile cases: 36 4.[b]Healthy subjects from endemic regions: 15
90%	145	160	<u>1.Cross-reactivity panel</u> <u>For antibody:</u> • Chikungunya IgM positive: 5 @ • Japanese Encephalitis IgM positive: 5 @ • West Nile Virus IgM positive: 5 * • Zika Virus IgM positive: 5 * <u>For antigen:</u> • Chikungunya PCR positive: 5 @ • Japanese Encephalitis NS1/PCR positive: 5 @ • West Nile Virus NS1/PCR positive: 5 * • Zika Virus NS1/PCR positive: 5 * 2.Rheumatoid Arthritis/other autoimmune disease cases: 30 3.[a]Acute febrile cases: 60 4.[b]Healthy subjects from endemic regions: 30
85%	206	220	<u>1.Cross-reactivity panel</u> <u>For antibody:</u> • Chikungunya IgM positive: 7 @ • Japanese Encephalitis IgM positive: 7 @ • West Nile Virus IgM positive: 7 * • Zika Virus IgM positive: 7 * <u>For antigen:</u> • Chikungunya PCR positive: 7 @ • Japanese Encephalitis NS1/PCR positive: 7 @ • West Nile Virus NS1/PCR positive: 7 *

			• Zika Virus NS1/PCR positive: 7 * 2.Rheumatoid Arthritis/other autoimmune disease cases: 34 3.[a]Acute febrile cases: 90 4.[b]Healthy subjects from endemic regions: 40
80%	258	260	1.Cross-reactivity panel For antibody: • Chikungunya IgM positive: 8 @ • Japanese Encephalitis IgM positive: 8 @ • West Nile Virus IgM positive: 8 * • Zika Virus IgM positive: 8 * For antigen: • Chikungunya PCR positive: 8 @ • Japanese Encephalitis NS1/PCR positive: 8 @ • West Nile Virus NS1/PCR positive: 8 * • Zika Virus NS1/PCR positive: 8 * 2.Rheumatoid Arthritis/other autoimmune disease cases: 45 3.[a]Acute febrile cases: 106 4.[b]Healthy subjects from endemic regions: 45

[a] Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR)
[b] Samples from healthy subjects from endemic regions negative for all Dengue markers (NS1, IgM, IgG, RNA)

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

Note: Commercially available validated standard panels that are accepted by accreditation agencies can also be used if the clinical sample with required analyte level is not available. If IgM positive samples for cross reactive flaviviruses are not available, commercially available IgM sera panel for different viruses can be procured and used to test cross reactivity.

Before used for evaluation, the NS1 reconstituted in serum samples needs to be tested by the reference assay and dilution which are positive only should be used for evaluation.

The serum sample used for spiking or reconstitution should be negative for Dengue NS1, RNA and IgM antibody.

6. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

7. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

8. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte(s) using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples for each analyte should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

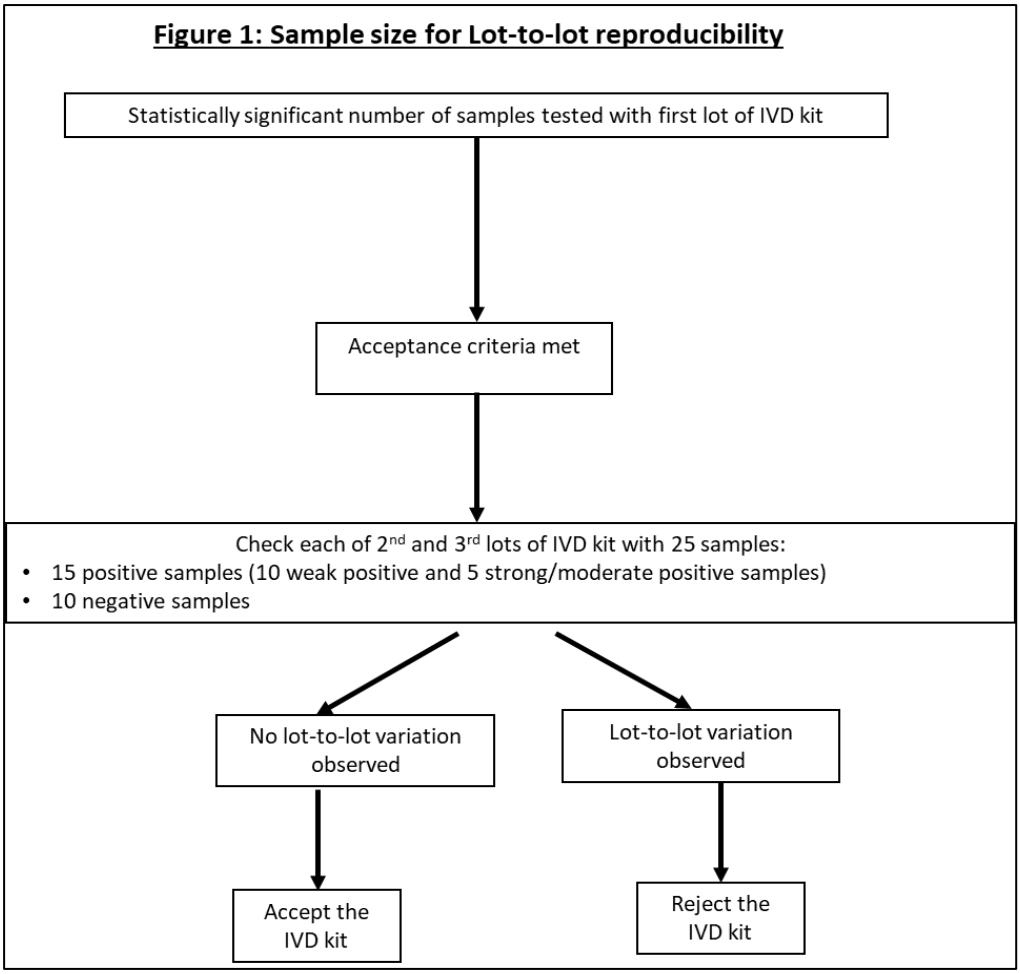
B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples per target analyte, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples per target analyte, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.



b. Inter-Operator variability: Testing and result interpretation should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples for each analyte should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples for each analyte should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples for each analyte should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

9. Interpretation of results:

Since the kits have been provided in combo format, concordance has to be calculated separately for NS1 and IgM, and the overall sensitivity and specificity have to be calculated based on the combined results of NS1 and IgM. If the sample is positive for any one or both analytes (NS1 or IgM or both), then the sample is considered positive. Refer the table below for interpretation:

NS1 Reference test result	IgM reference test result	Final Reference test result	NS1 Index test result	IgM Index test result	Final index test result	Interpretation for dengue positivity
+	+	Positive	+	-	Positive	True Positive
+	+	Positive	-	+	Positive	True Positive
+	+	Positive	-	-	Negative	False Negative
+	+	Positive	+	+	Positive	True Positive
+	-	Positive	+	-	Positive	True Positive
+	-	Positive	-	+	Positive	True Positive
+	-	Positive	-	-	Negative	False Negative
-	+	Positive	+	-	Positive	True Positive
-	+	Positive	-	+	Positive	True Positive
-	+	Positive	-	-	Negative	False Negative
-	-	Negative	-	+	Positive	False Positive
-	-	Negative	+	-	Positive	False Positive

Consider one-on-one results for analyte-specific analysis.

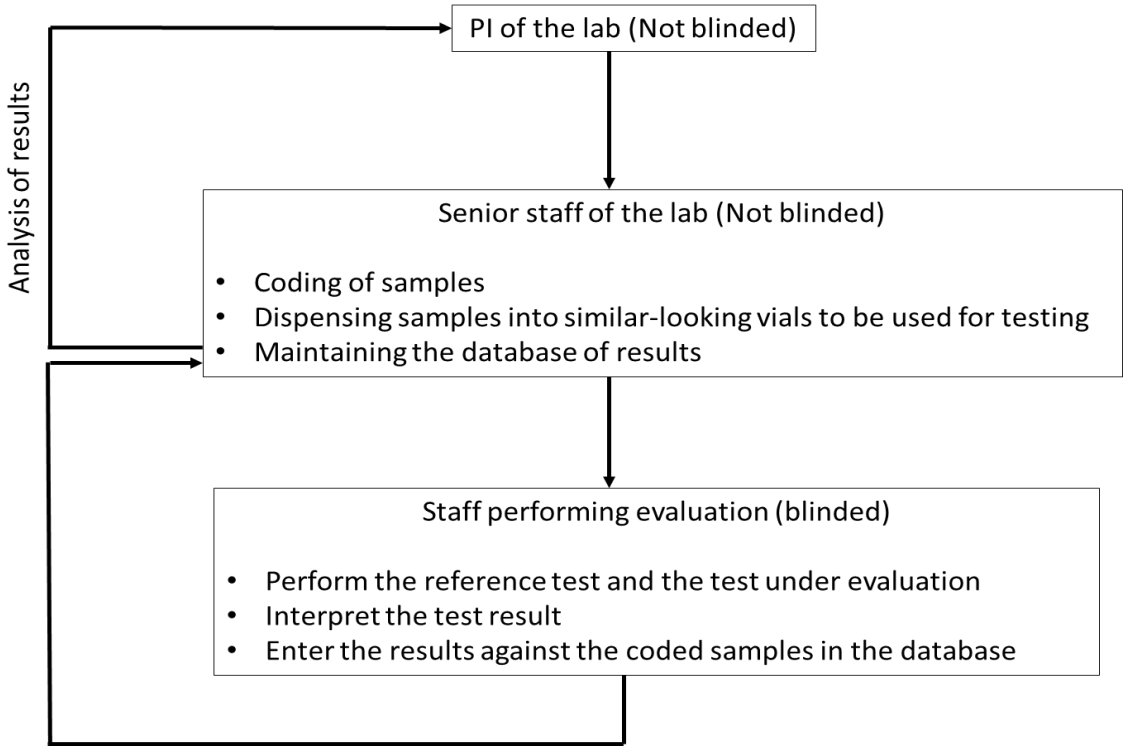
10. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

11. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



12. Acceptance criteria:

A minimum concordance of 80% for NS1 and 80% for IgM should be achieved with the reference assay, and an overall combined sensitivity***** and specificity**$** of ≥90% each.

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

Invalid test rate: ≤5%

***** Samples which are positive for NS1 or IgM or both by the kit under evaluation (irrespective of the reference assay results) will be considered as positive for dengue and used for overall sensitivity calculation. To achieve at least the overall combined performance characteristics outlined in the acceptance criteria, ≥160 positive samples and ≥160 negative samples should be used for evaluation.

$ Sample which are negative for both NS1 and IgM by kit under evaluation (irrespective of the reference assay results) will be considered negative for dengue and used for overall specificity calculation

13. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, Vazquez S, Cartozian E, Pelegrino JL, Artsob H, Guzman MG, Olliaro P, Zwang J, Guillerm M, Kliks S, Halstead S, Peeling RW, Margolis HS. Evaluation of commercially available diagnostic tests for the detection of Dengue virus NS1 antigen and anti-Dengue virus IgM antibody. PLoSNegl Trop Dis. 2014 Oct 16;8(10):e3171. doi: 10.1371/journal.pntd.0003171.
2. Hermann LL, Thaisomboonsuk B, Poolpanichupatam Y, Jarman RG, Kalayanarooj S, Nisalak A, Yoon IK, Fernandez S. Evaluation of a Dengue NS1 antigen detection assay sensitivity and specificity for the diagnosis of acute Dengue virus infection. PLoSNegl Trop Dis. 2014 Oct 2;8(10):e3193. doi: 10.1371/journal.pntd.0003193.
3. Yow KS, Aik J, Tan EY, Ng LC, Lai YL. Rapid diagnostic tests for the detection of recent Dengue infections: An evaluation of six kits on clinical specimens. PLoS One. 2021 Apr 1;16(4):e0249602. doi: 10.1371/journal.pone.0249602.
4. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
5. WHO, Evaluation of commercially available anti-Dengue virus immunoglobulin M tests. (Diagnostics evaluation series, 3). ISBN 978 92 4 159775 3.
6. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE NS1 and IgM COMBO RDT KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak/simulated samples)	
	Negative samples (provide details: clinical/spiked, including cross reactivity panel/simulated samples)	

Results:

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Dengue NS1 and IgM combo RDT kit	Positive			
	Negative			
	Total			

Prepare the above 2x2 table individually for each analyte, as well as for overall performance characteristics

	Estimate (%)	95% CI
Combined Sensitivity		
Combined Specificity		
Dengue NS1 standalone sensitivity		
Dengue NS1 standalone specificity		
Dengue IgM standalone sensitivity		
Dengue IgM standalone specificity		

- Cross-reactivity:
- Invalid test rate:
  - **Conclusions:**
    - Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ….. sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for …….. Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……

Field evaluation protocol for Dengue NS1 and IgM combo RDT kits

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Dengue NS1/IgM RDT combo kits in the diagnosis of Dengue infection in individuals with unknown disease status.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Laboratory supplies

IV. Ethical approval:

The study will be initiated after approval from the institutional human ethics committee.

V. Procedure:

1. **Study design/type**: Cross-sectional study
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through
   A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel
   - Handling of Dengue NS1 RDT/IgM RDT IVD kits received for performance evaluation (Verification/Storage/Unpacking etc).
   - Testing, interpreting, recording of results & reporting
   - Data handling, data safety & confidentiality

3. Sample size for performance evaluation:

Sample sizes of positive and negative samples of Dengue against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. It is further assumed that 30% of the individuals attending the health care facilities for acute febrile illness and suspected for Dengue will be positive for Dengue. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x P x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x P x } (1 - \text{IR})}$$

- *n (se) is the minimum number of number of individuals to be enrolled to obtain the requisite number of positive samples.*
- *n (sp) is the minimum number of number of individuals to be enrolled to obtain the requisite number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *P is the prevalence of the disease*
- *d is the predetermined marginal error (5%).*
- *IR is the invalid test rate*

Sample size has to be calculated based on both the sensitivity and the specificity. The final sample size will be the maximum of the two. For example, at 95% sensitivity and 95% specificity, the sample size required will be 260 (maximum of 260 and 110). It is desirable to cover at least one Dengue season so that adequate samples are available for evaluation.

Table 1. Sample sizes for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of individuals to be enrolled[*] [Sample size rounded off] #*
99%	53	60
95%	255	260
90%	484	490
85%	686	690
80%	861	870
[*] Individuals attending the health care facilities for acute febrile illness and suspected for Dengue meeting the inclusion criteria		

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of individuals to be enrolled[*] [Sample size rounded off] #*
99%[#]	23	30
95%	109	110
90%	207	210
85%	294	300
80%	369	370
[*] Individuals attending the health care facilities for acute febrile illness and suspected for Dengue meeting the inclusion criteria		

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Recruitment of cases shall be halted once desired number of positive and negative samples are reached.

4. Inclusion criteria:

Patient with Dengue like illness (A patient with acute febrile illness of 1-14 days with two or more manifestations: Head ache, retro-orbital pain, myalgia, arthralgia, rash, hemorrhagic manifestations etc. The 1-14 days disease duration shall cover viraemic as well as convalescent phase of Dengue infection, so that both Dengue NS1 and IgM positive cases are enrolled.)

5. Exclusion criteria:

Individuals with already known positive history for other pathogens

6. Reference assay:

Anti-DENV IgM detection ELISA WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved kit

AND/OR

DENV NS1 ELISA WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved kit

Serotype status to be assessed using a combination of CDC and/or NIV real-time PCR serotyping protocols.

7. Study implementation:

The individuals with Dengue like illness will be recruited into the study and five ml of whole blood will be collected in vacutainer tubes and the serum will be separated by centrifugation and used for the study.

It needs to be ensured that the samples are tested by reference tests and index test simultaneously.

8. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

9. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

10. Positive samples:

Samples which are positive for IgM or NS1 or both by the reference assays will be considered as true positive samples.

11. Negative samples:

Samples which are negative by the reference assay will be considered as negative.

N.B.:

False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

A. Cross reactivity (other flavivirus infections):
A.1 NS1:

Clinical samples or commercially available NS1 antigens from other flaviviruses will be used to test cross reactivity of the NS1 component of index test.

i. Chikungunya PCR positive: 5 samples @
ii. Japanese Encephalitis PCR/antigen positive: 5 samples @
iii. West Nile Virus PCR/antigen: 5 samples *
iv. Zika Virus PCR/antigen: 5 samples *

Cross reactivity panel is arranged in descending order of priority.

The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

In the absence of natural samples, spiked samples may be used, as per details provided in the note below.

Note:

Commercially available validated standard panels that are accepted by accreditation agencies can also be used if the clinical sample with required analyte level is not available. Recombinant NS1 antigen of cross reactive flaviviruses (Zika, West Nile and Japanese Encephalitis viruses) expressed in mammalian cells can be obtained commercially and reconstituted in serum samples (100 ng -1 µg/ml) and diluted in the ratio of 1:2 and used accordingly (at least five dilutions for each virus specific NS1).

Before used for evaluation, NS1 reconstituted in serum samples needs to be tested by the reference assay and dilution which are positive only should be used for evaluation.

The serum samples used for reconstitution should be negative for Dengue NS1, RNA and IgM antibody.

A.2 IgM:
Clinical samples positive for IgM for other flaviviruses will be used to test cross reactivity of the IgM component of index test.

i. Chikungunya IgM positive: 5 samples @
ii. Japanese Encephalitis IgM positive: 5 samples @
iii. West Nile Virus IgM positive: 5 samples *
iv. Zika Virus IgM positive: 5 samples *

Cross reactivity panel is arranged in descending order of priority.

The pathogens marked @ are essentially to be tested.

*It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

Note: Depending on the availability of IgM positive samples for cross reactive flaviviruses, the requirement of samples for each virus may be increased or decreased accordingly to reach the total number of samples. If IgM positive samples for cross reactive flaviviruses are not available, commercially available IgM sera panel for different viruses can be procured and used to test cross reactivity.

12. Statistical analysis:

Concordance will be calculated separately for Dengue NS1 and IgM. Combined sensitivity and specificity will also be calculated.

Interim analysis of data shall be conducted on completing evaluation of 25%, 50% and 75% of samples. If, at any point, the performance of the assay is found to be not satisfactory, the assay shall not be evaluated further. Evaluation fee shall be charged accordingly.

13. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte(s) using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples for each analyte should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

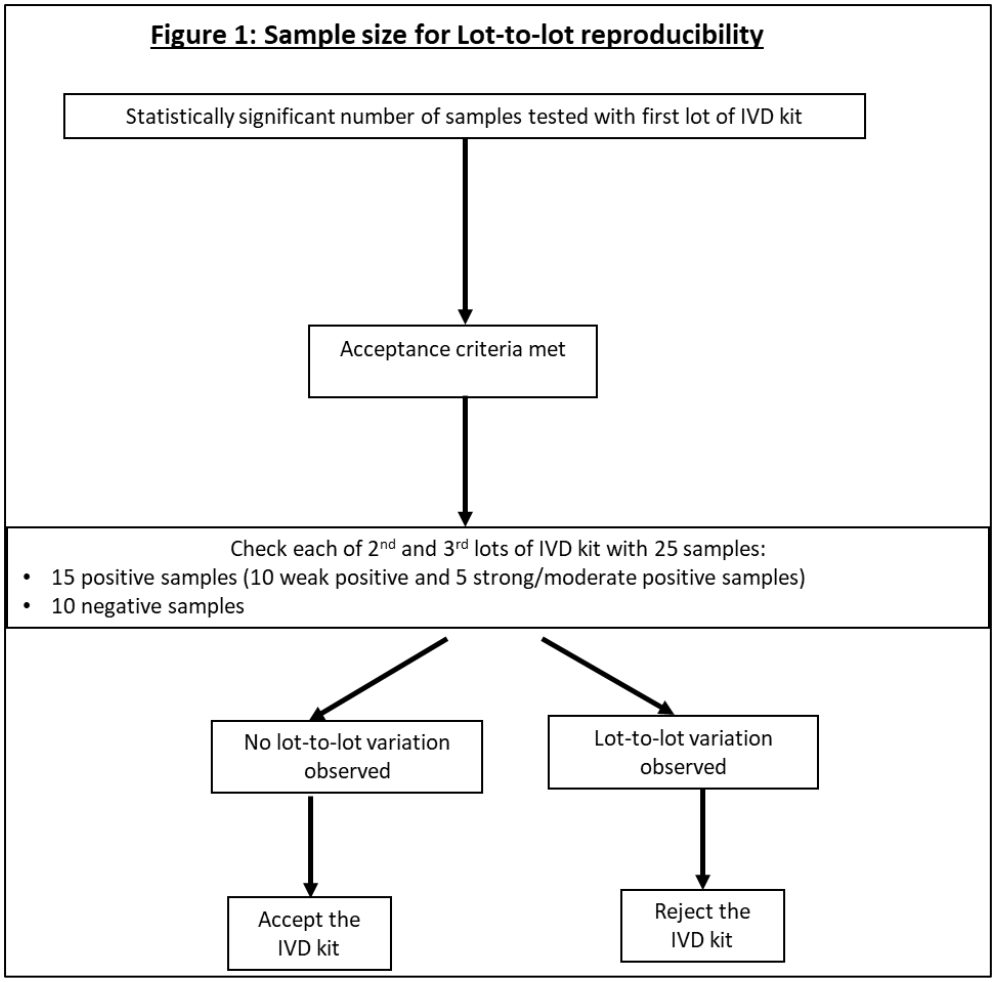
B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples per target analyte, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples per target analyte, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.



b. Inter-Operator variability: Testing and result interpretation should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples for each analyte should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples for each analyte should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples for each analyte should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

14. Interpretation of results:

Since the kits have been provided in a combo format, the sensitivity and specificity has to be calculated based on the combined results of the NS1 and IgM. If the sample is positive for any one or both analytes (NS1 or IgM or both), then the sample is considered positive. Refer the table below:

NS1 Reference test result	IgM reference test result	Final Reference test result	NS1 Index test result	IgM Index test result	Final index test result	Interpretation for dengue positivity
+	+	Positive	+	-	Positive	True Positive
+	+	Positive	-	+	Positive	True Positive
+	+	Positive	-	-	Negative	False Negative
+	+	Positive	+	+	Positive	True Positive
+	-	Positive	+	-	Positive	True Positive
+	-	Positive	-	+	Positive	True Positive
+	-	Positive	-	-	Negative	False Negative
-	+	Positive	+	-	Positive	True Positive
-	+	Positive	-	+	Positive	True Positive
-	+	Positive	-	-	Negative	False Negative
-	-	Negative	-	+	Positive	False Positive
-	-	Negative	+	-	Positive	False Positive

Consider one-on-one results for analyte-specific analysis.

15. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

16. Blinding of laboratory staff

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the results of the reference assay. The PI of the evaluation exercise should remain unblinded, i.e., privy to the results of the reference test. Another senior laboratory staff selected by the PI may remain unblinded for overseeing the activity and maintaining the database of results, and the PI should maintain confidentiality of data. The data should be analyzed only by the PI of the evaluating lab.

17. Acceptance criteria:

A minimum concordance of 80% for NS1 and 80% for IgM should be achieved with the reference assay, and an overall combined sensitivity***** and specificity**$** of ≥90% each.

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

Invalid test rate: ≤5%

***** Samples which are positive for NS1 or IgM or both by the kit under evaluation (index test) irrespective of the reference assay results will be considered positive for dengue and used for overall sensitivity calculation

$ Samples which are negative for both NS1 and IgM by kit under evaluation will be considered as negative for dengue and used for overall specificity calculation

To achieve at least the overall performance characteristics outlined in the acceptance criteria (≥90% overall sensitivity and ≥90% overall specificity), ≥490 individuals satisfying the case definition need to be enrolled to obtain the requisite number of positive samples. This sample size is sufficient for required number of negative samples.

Recruitment should be terminated once the desired number of positive cases is enrolled and tested.

18. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, Vazquez S, Cartozian E, Pelegrino JL, Artsob H, Guzman MG, Olliaro P, Zwang J, Guillerm M, Kliks S, Halstead S, Peeling RW, Margolis HS. Evaluation of commercially available diagnostic tests for the detection of Dengue virus NS1 antigen and anti-Dengue virus IgM antibody. PLoSNegl Trop Dis. 2014 Oct 16;8(10):e3171. doi: 10.1371/journal.pntd.0003171.
2. Hermann LL, Thaisomboonsuk B, Poolpanichupatam Y, Jarman RG, Kalayanarooj S, Nisalak A, Yoon IK, Fernandez S. Evaluation of a Dengue NS1 antigen detection assay sensitivity and specificity for the diagnosis of acute Dengue virus infection. PLoSNegl Trop Dis. 2014 Oct 2;8(10):e3193. doi: 10.1371/journal.pntd.0003193.
3. Ganeshkumar P, Murhekar MV, Poornima V, Saravanakumar V, Sukumaran K, Anandaselvasankar A, John D, Mehendale SM. Dengue infection in India: A systematic review and meta-analysis. PLoSNegl Trop Dis. 2018 Jul 16;12(7):e0006618. doi: 10.1371/journal.pntd.0006618.
4. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1

5. WHO, Evaluation of commercially available anti-Dengue virus immunoglobulin M tests. (Diagnostics evaluation series, 3). ISBN 978 92 4 159775 3.
6. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE NS1 and IgM COMBO RDT KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples: Not applicable, may categorize cases as per duration of illness	
	Negative samples (may categorize as per duration of illness, must include cross reactivity panel)	

Results

		Reference assay …….…… (name)		
		Positive	Negative	Total
Name of NS1 and IgM combo RDT kit	Positive			
	Negative			
	Total			

Prepare the above 2x2 table individually for each analyte, as well as for overall performance characteristics

	Estimate (%)	95% CI
Combined Sensitivity		
Combined Specificity		
Denue NS1 standalone sensitivity		
Dengue NS1 standalone specificity		
Dengue IgM standalone sensitivity		
Dengue IgM standalone specificity		

- Cross-reactivity:
- Invalid test rate:
  - Conclusions:
    - Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in using kits provided by the manufacturer from the batch mentioned above using ….. sample in ……. (field/controlled lab). Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for ……..NS1 and IgM combo Kit (Lot No……) manufactured by …… (supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In charge …… Seal ……

Performance evaluation protocol for Dengue real-time PCR kit

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

This recommendation focuses on the laboratory performance evaluation of Dengue virus molecular diagnostic test. All clinical samples tested in the study should be evaluated in accordance with the candidate test's instructions for use.

II. Purpose:

To evaluate the performance characteristics of Dengue real-time PCR kits in the diagnosis of Dengue infection using irreversibly de-identified leftover archived/ spiked clinical samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.
2. **Preparation of Evaluation sites/laboratories:**

Identified IVD kit evaluation laboratories should establish their proficiency through

A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

- Preparation & characterization of kit evaluation panel
- Handling of Dengue RT-PCR kits received for performance evaluation (Verification/Storage/Unpacking etc).
- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Dengue RNA evaluation panel:

Well characterised Dengue serum/plasma panel positive for RNA by RT-PCR is a critical requirement for performance evaluation of IVD kits utilizing genome detection. Hence statistically significant number of sera/plasma samples should be available from Dengue PCR confirmed cases.

4. RNA extraction:

RNA extraction should be performed as per reference test and index test Instructions for Use (IFU).

If the manufacturer of the index test recommends a specific RNA extraction kit/system, the same needs to be provided by the manufacturer if the evaluating lab is unable to procure the same/it is not available within the lab's IVD evaluation scope.

5. Real-Time PCR System:

PCR should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Real-time closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.

6. Internal control/Extraction control:

The test under evaluation should have an internal control or extraction control (RNA added before extraction to a sample).

7. Reference assay:

Any WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Dengue PCR assay or CDC/NIV protocol for detection of Dengue virus RNA should be used as the reference assay.

All positive samples should be confirmed positive for at least one serotype by real-time PCR assay using CDC/NIV protocol.

All negative samples should be negative for all the markers of Dengue infection (NS1, IgM, and RNA).

8. **Sample size and sample panel composition:** Sample sizes of positive and negative samples and sample panel composition against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Positive samples: These include samples positive by the reference real-time PCR assay (True positives) and representative of all four serotypes.

Negative samples: All negative samples should be negative by reference real-time PCR assay, WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved NS1 antigen ELISA kit-and WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved IgM Capture ELISA.

Table 1. Sample sizes and panel composition of positive Dengue samples for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%	16	20	Strong positive (Ct value <25): 6 Moderate positive (Ct value between 25-30): 7 Weak positive (Ct value >30 to 34): 7
95%	77	80	Strong positive (Ct value <25): 24 Moderate positive (Ct value between 25-30): 28 Weak positive (Ct value >30 to 34): 28
90%	145	150	Strong positive (Ct value <25): 44 Moderate positive (Ct value between 25-30): 53 Weak positive (Ct value >30 to 34): 53
85%	206	210	Strong positive (Ct value <25): 62 Moderate positive (Ct value between 25-30): 74 Weak positive (Ct value >30 to 34): 74
80%	258	260	Strong positive (Ct value <25): 78 Moderate positive (Ct value between 25-30): 91 Weak positive (Ct value >30 to 34): 91
Note: If clinical samples positive for a particular serotype is not available, tissue culture fluid (5-10 different isolates with a plaque forming unit of 10^{5-6}/ml) (Heat-inactivated) from reference laboratories can be obtained, spiked in serum samples (15 µl isolate + 150 µl) and can be further diluted in the ratio of 1:10, frozen at -80°C, and tested by the reference assay when needed and the positive samples can be used for evaluation. The serum used for spiking isolate should be negative for Dengue virus RNA, and NS1.			

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes and panel composition of negative Dengue samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%	16	20	1.Cross-reactivity panel: • Chikungunya PCR positive: 4 @ • Japanese Encephalitis PCR positive: 1 @ • Zika Virus PCR positive: 1 * • West Nile Virus PCR positive: 1 * 2.[a]Acute febrile cases: 8 3.[b]Healthy subjects from endemic regions: 5
95%	77	80	1.Cross-reactivity panel: • Chikungunya PCR positive: 15 @ • Japanese PCR Encephalitis positive: 5 @ • Zika Virus PCR positive: 5 * • West Nile Virus PCR positive: 5 * 2.[a]Acute febrile cases: 30 3.[b]Healthy subjects from endemic regions: 20
90%	145	150	1.Cross-reactivity panel: • Chikungunya PCR positive: 28 @ • Japanese Encephalitis PCR positive: 9 @ • Zika Virus PCR positive: 9 * • West Nile Virus PCR positive: 9 * 2.[a]Acute febrile cases: 57 3.[b]Healthy subjects from endemic regions: 38
85%	206	210	1.Cross-reactivity panel: • Chikungunya PCR positive: 39 @ • Japanese Encephalitis PCR positive: 13 @ • Zika Virus PCR positive: 13 * • West Nile Virus PCR positive: 13 * 2.[a]Acute febrile cases: 79 3.[b]Healthy subjects from endemic regions: 53
80%	258	260	1.Cross-reactivity panel: • Chikungunya PCR positive: 49 @ • Japanese Encephalitis PCR positive: 16 @

			• Zika Virus PCR positive: 16 * • West Nile Virus PCR positive: 16 * 2.[a]Acute febrile cases: 98 3.[b]Healthy subjects from endemic regions: 65

[a] Acute febrile cases negative for all markers of Dengue (NS1 & IgM & IgG & RNA)
[b] Samples from healthy subjects from endemic regions negative for all Dengue markers (NS1, IgM, IgG, nucleic acid)

* Note:
Commercially available validated standard panels that are accepted by accreditation agencies can also be used if the clinical sample with required analyte level is not available. If PCR positive samples for cross reactive flaviviruses not available, commercially available RNA panels should be used to test cross reactivity.

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

9. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

10. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

11. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

12. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

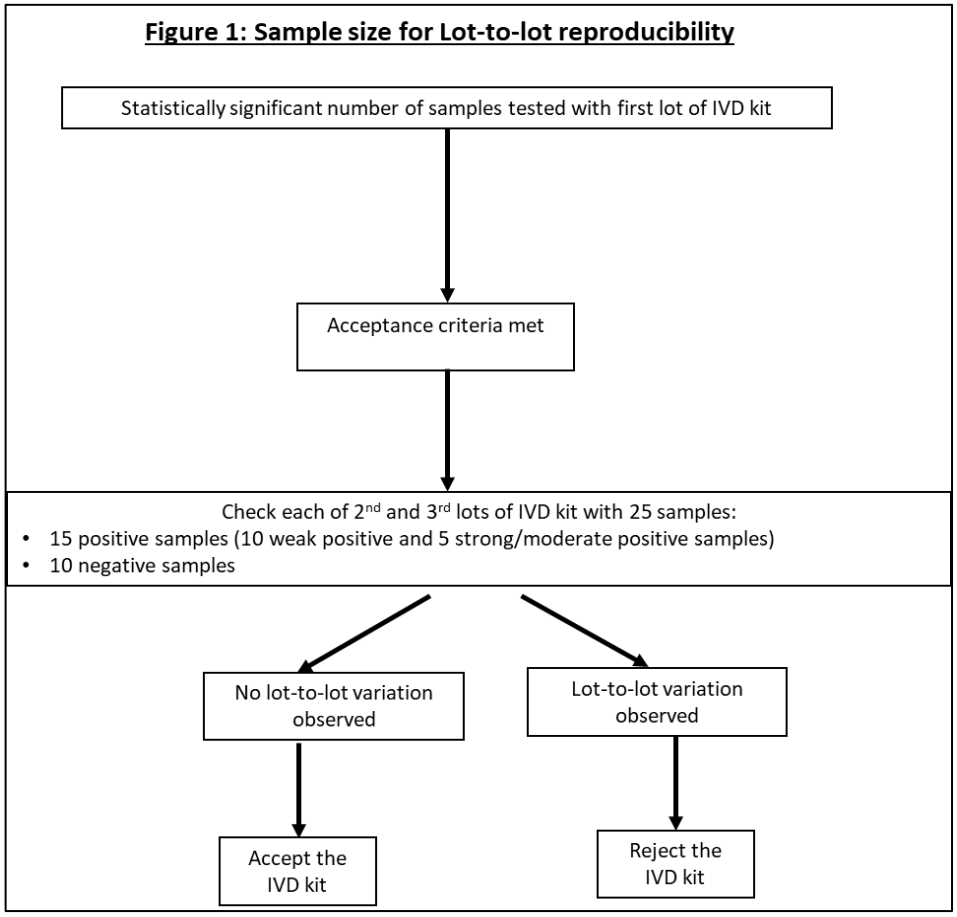
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

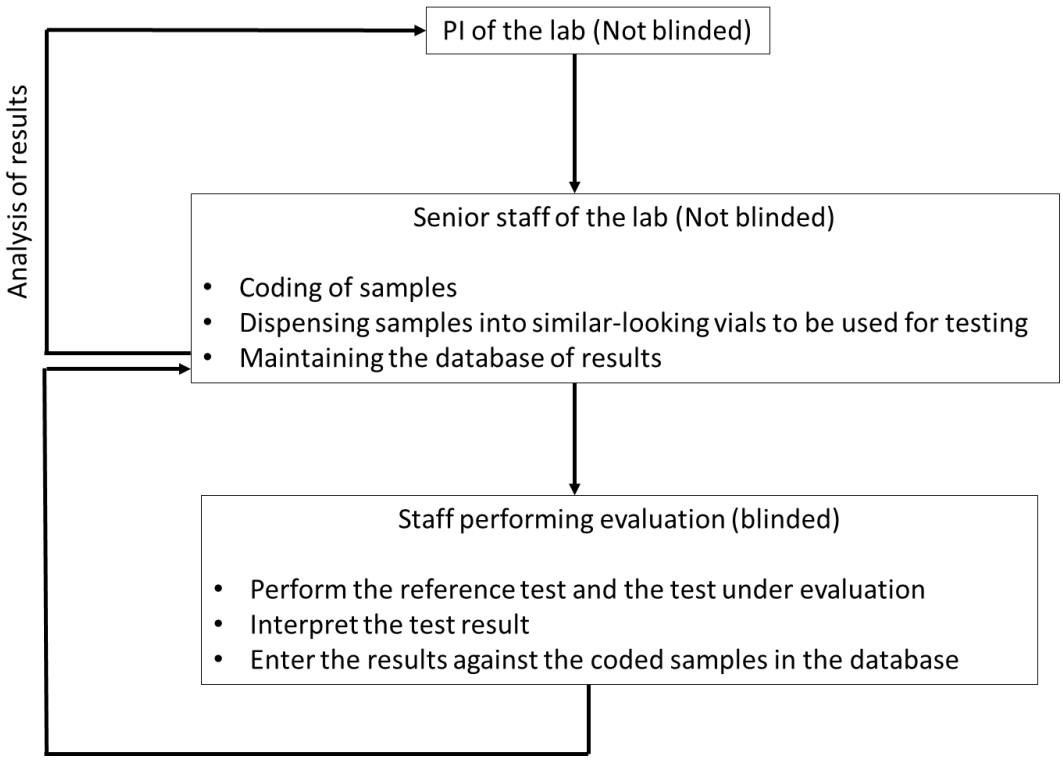
d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report Ct standard deviation across repeats (≤1cycle deviation recommended).

13. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



14. Acceptance Criteria:

Sensitivity: ≥95%

Specificity: ≥98%

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥80 positive samples and ≥80 negative samples should be used for evaluation.

15. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered

only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Santiago, G.A., Vázquez, J., Courtney, S. et al. Performance of the Trioplex real-time RT-PCR assay for detection of Zika, Dengue, and Chikungunya viruses. Nat Commun 9, 1391 (2018). https://doi.org/10.1038/s41467-018-03772-1
2. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
3. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE REAL-TIME PCR KITS

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: clinical/ spiked, strong, moderate, weak/simulated samples)	
	Negative samples (provide details clinical/ spiked, including cross reactivity panel/simulated samples)	

Results

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Dengue real-time PCR kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:
- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ..... sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for Dengue…….. Kit (Lot No……) manufactured by …… (supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal …….

Field evaluation protocol for Dengue real-time PCR kits

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Dengue real-time PCR kits in the diagnosis of Dengue infection in individuals with unknown disease status.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Laboratory supplies

IV. Ethical approvals:

The study will be initiated after approval from the institutional human ethics committee.

V. Procedure:

1. **Study design/type**: Cross-sectional study
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through
   A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel
   - Handling of Dengue RT-PCR kits received for performance evaluation (Verification/Storage/Unpacking etc).
   - Testing, interpreting, recording of results & reporting
   - Data handling, data safety & confidentiality

3. **Sample size for performance evaluation**: Sample sizes of positive and negative samples of Dengue against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. It is further assumed that 30% of the

individuals attending the health care facilities for acute febrile illness and suspected for Dengue will be positive for Dengue. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x P x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x P x } (1 - \text{IR})}$$

- *n (se) is the minimum number of number of individuals to be enrolled to obtain the requisite number of positive samples.*
- *n (sp) is the minimum number of number of individuals to be enrolled to obtain the requisite number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *P is the prevalence of the disease*
- *d is the predetermined marginal error (5%).*
- *IR is the invalid test rate*

Sample size has to be determined based on both the sensitivity and the specificity. The required sample size will be the maximum of the two. For example, at 95% sensitivity and 95% specificity, the sample size required will be 260 (maximum of 260 and 110).

Table 1. Sample sizes for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of individuals to be enrolled* [Sample size rounded off] #*

99%	53	60
95%	255	260
90%	484	490
85%	686	690
80%	861	870
* Individuals attending the health care facilities for acute febrile illness and suspected for Dengue meeting the inclusion criteria		

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of individuals to be enrolled** *[Sample size rounded off] #*
99%	23	30
95%	109	110
90%	207	210
85%	294	300
80%	369	370
* Individuals attending the health care facilities for acute febrile illness and suspected for Dengue meeting the inclusion criteria		

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Recruitment of cases shall be halted once desired number of positive and negative samples are reached.

4. Inclusion criteria:

Individuals with Dengue like illness (A patient with acute febrile illness of 2-7 days with two or more manifestations: Head ache, retro-orbital pain, myalgia, arthralgia, rash, hemorrhagic manifestations)

5. Exclusion criteria:

Individuals with already known positive history for other pathogens

6. RNA extraction

RNA extraction should be performed as per reference test and index test Instructions for Use (IFU).

If the manufacturer of the index test recommends a specific RNA extraction kit/system, the same needs to be provided by the manufacturer if the evaluating lab is unable to procure the same/it is not available within the lab's IVD evaluation scope.

7. Real-Time PCR System:

PCR should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Real-time closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.

8. Internal control/Extraction control:

The test under evaluation should have an internal control or extraction control (RNA added before extraction to a sample).

9. Reference assay:

Any WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Dengue PCR assay or CDC/NIV protocol for detection of Dengue RNA should be used as the reference assay.

All positive samples should be confirmed positive for at least one serotype by real-time PCR assay using CDC/NIV protocol.

All negative samples should be negative for all the markers of Dengue infection (NS1 & IgM & IgG and RNA).

10. Study implementation:

The individuals with Dengue like illness will be recruited into the study and five ml of whole blood will be collected in vacutainer tubes and the serum will be separated by centrifugation and used for the study.

It needs to be ensured that the samples are tested by reference tests and index test simultaneously.

11. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

12. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

13. Positive samples:

Samples which are positive by reference real-time PCR assay will be considered as true positive sample.

14. Negative samples:

Samples which are negative by the reference assay will be considered as negative.

N.B.:

False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

A. Cross reactivity:

Clinical samples or commercially available validated viral RNA genome of other flaviviruses that are accepted by accreditation agencies/RNA from sequence confirmed virus isolates will be used to test cross reactivity of the index test.

a. Chikungunya PCR positive: 5 samples @
b. Japanese Encephalitis PCR positive: 5 samples @
c. West Nile Virus PCR positive: 5 samples *
d. Zika Virus PCR positive: 5 samples *

Alternatively, tissue culture fluid of cross reactive flaviviruses (with a plaque forming unit of 10^{5-6}/ml)(Heat inactivated) from reference laboratories can be obtained, spiked in serum samples (15 µl isolate + 150 µl) and can be further diluted in the ratio of 1:10, tested by the reference assay and the negative samples can be used for evaluation.

The serum used for spiking isolate should be negative for Dengue virus RNA, and NS1.

Cross reactivity panel is arranged in descending order of priority.

The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

15. Statistical analysis:

Sensitivity and specificity will be calculated.

Interim analysis of data shall be conducted on completing evaluation of 25%, 50% and 75% of samples. If, at any point, the performance of the assay is found to be not satisfactory, the assay shall not be evaluated further. Evaluation fee shall be charged accordingly.

16. Repeatability and Reproducibility Assessment:

A. <u>Repeatability Assessment</u>

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

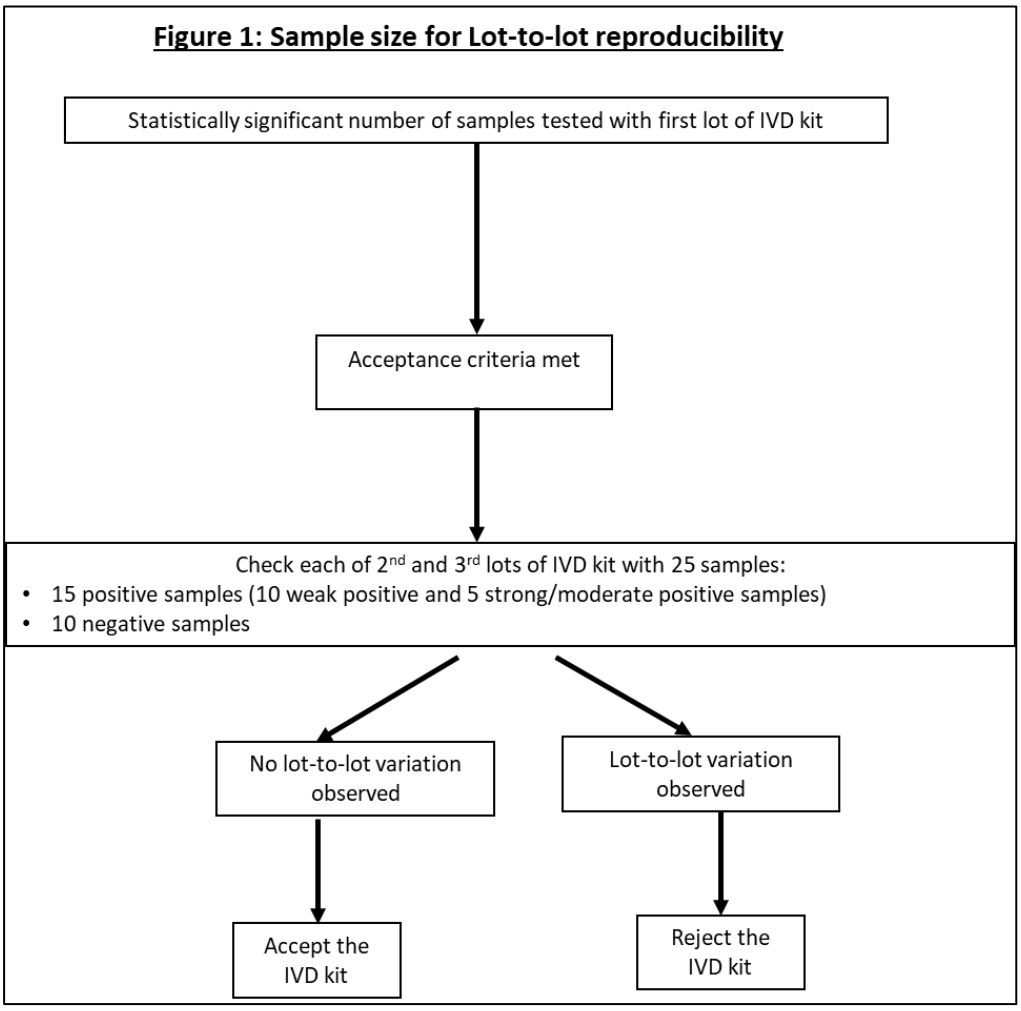
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report Ct standard deviation across repeats (≤1cycle deviation recommended).

17. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

18. Blinding of laboratory staff

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the results of the reference assay. The PI of the evaluation exercise should remain unblinded, i.e., privy to the results of the reference test. Another senior laboratory staff selected by the PI may remain unblinded for overseeing the activity and maintaining the database of results, and the PI should maintain confidentiality of data. The data should be analyzed only by the PI of the evaluating lab.

19. Acceptance Criteria

Sensitivity: ≥95%

Specificity: ≥98%

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥260 individuals satisfying the case definition need to be enrolled to obtain the requisite number of positive samples. This sample size is sufficient for required number of negative samples.

Recruitment should be terminated once the desired number of positive cases is enrolled and tested.

20. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Santiago, G.A., Vázquez, J., Courtney, S. et al. Performance of the Trioplex real-time RT-PCR assay for detection of Zika, Dengue, and Chikungunya viruses. Nat Commun 9, 1391(2018). https://doi.org/10.1038/s41467-018-03772-1
2. Ganeshkumar P, Murhekar MV, Poornima V, Saravanakumar V, Sukumaran K, Anandaselvasankar A, John D, Mehendale SM. Dengue infection in India: A systematic review and meta-analysis. PLoSNegl Trop Dis. 2018 Jul 16;12(7):e0006618. doi: 10.1371/journal.pntd.0006618.
3. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
4. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE REAL-TIME PCR KITS

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples: Not applicable, may categorize cases as per duration of illness	
	Negative samples (may categorize as per duration of illness, must include cross reactivity panel)	

Results

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Dengue real-time PCR kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:
- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ..... sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for Dengue…….. Kit (Lot No……) manufactured by …… (supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……..

Performance evaluation protocol for Real-time PCR tests for Zika virus

I. Background:

CDSCO and ICMR, New Delhi, aimed at facilitating the evaluation and deployment of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding in-vitro diagnostic kit (IVD) performance.

This recommendation focuses on the laboratory performance evaluation of Zika virus molecular diagnostic test. All clinical samples tested in the study should be evaluated in accordance with the candidate test's proposed diagnostic algorithm (i.e., tested using the procedure in the instructions for use), including retesting when appropriate.

II. Purpose:

To evaluate the performance characteristics of Zika virus RT-PCR test for diagnosis of Zika infection using irreversibly de-identified leftover archived/ spiked clinical samples.

III. Requirements:

1. Supply of kits under evaluation (along with batch/lot No. Expiry & required details)
2. Evaluation site/laboratory should be equipped with necessary equipment and supplies for molecular testing. Any essential equipment and consumables for closed system to be supplied and maintained from the manufacturer, during the period of evaluation.
3. Reference test kits
4. Characterized evaluation panel
5. Laboratory supplies

IV. Ethics approval:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.
Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.

2. **Preparation of Evaluation site/laboratory:** Performance evaluation performance and report to be issued only from designated reference testing laboratory/ ISO accredited laboratory, as specified by state or central licensing authority.

3. **Identified IVD kit evaluation laboratories should establish their proficiency through:**

A. Accreditation for at least one Quality management system (accreditation for Testing Laboratory/ Calibration Laboratory (ISO/IEC 17025), Medical Laboratory (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

B. Staff training: All the staff involved in the IVD kit evaluation should undergo hands on training and competency testing on following

- Preparation & characterization of evaluation panel
- Handling of Zika molecular diagnostic kits received for performance evaluation (Verification/Storage/Unpacking etc.)
- Testing, interpretation, recording of results & reporting
- Data handling, data safety & confidentiality

4. Preparation of Zika reference evaluation panel:

Well characterized Zika molecular evaluation panel is a critical requirement for performance evaluation of IVD kits. Hence, statistically significant number of clinical samples should be used for evaluation.

- Frozen samples (≤-70°C) may be used, if stored appropriately and analytical data demonstrate that accuracy of test results is not affected.
- Samples that previously tested positive by WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved PCR and/or CDC/NIV approved protocols may be used.
- In the absence of natural samples, spiked clinical samples may be used.

5. RNA extraction:

RNA extraction should be performed as per reference test and index test Instructions for Use (IFU).

If the manufacturer of the index test recommends a specific RNA extraction kit/system, the same needs to be provided by the manufacturer if the evaluating lab is unable to procure the same/it is not available within the lab's IVD evaluation scope.

6. Real-Time PCR System:

PCR should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Real-time closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.

7. Internal control/Extraction control:

The test under evaluation should have an internal control or extraction control (RNA added before extraction to a sample).

8. Reference assay:

Any WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Zika PCR assay or CDC/NIV protocol for detection of Zika RNA should be used as the reference assay.

Evaluations with the reference test should be conducted as per the manufacturer's instructions for use.

Positive and negative samples should be subjected to both the reference test and test under evaluation.

9. Sample size and sample panel composition: Sample sizes of positive and negative samples and panel composition against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Positive Samples:

- Clinical positive samples: Sample tested positive by Zika virus molecular reference assay from clinically suspect cases.
- Contrived positive samples: In absence of reference clinical samples, a contrived positive sample may be used.

Contrived positive samples should be prepared using spiking of diluted Zika virus culture isolate in unique negative samples, as per the note below:

Table 1. Sample sizes and panel composition of positive Zika virus samples for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%[#]	16	20	Strong positive (Ct value <25): 6 Moderate positive (Ct value between 25-30): 7 Weak positive (Ct value >30 to 34): 7
95%	77	80	Strong positive (Ct value <25): 24 Moderate positive (Ct value between 25-30): 28 Weak positive (Ct value >30 to 34): 28
90%	145	150	Strong positive (Ct value <25): 44 Moderate positive (Ct value between 25-30): 53 Weak positive (Ct value >30 to 34): 53
85%	206	210	Strong positive (Ct value <25): 62 Moderate positive (Ct value between 25-30): 74 Weak positive (Ct value >30 to 34): 74
80%	258	260	Strong positive (Ct value <25): 78 Moderate positive (Ct value between 25-30): 91 Weak positive (Ct value >30 to 34): 91

Note 1: Representative positive samples from genotype (African, Asian/American) may be included, if feasible.

Note 2: Contrived positive samples – In absence of reference clinical samples, a contrived positive sample may be used.

Contrived positive samples should be prepared using spiking of diluted Zika virus culture isolate in unique negative samples, as follows:

Tissue culture fluid (3-5 different isolates with a plaque forming unit of 10^{5-6}/ml) (Heat inactivated) from reference laboratories can be obtained, spiked in serum samples (15 µl isolate + 150 µl) and can be further diluted in the ratio of 1:10, tested by the reference assay and the positive samples can be used for evaluation.
The serum used for spiking isolate should be negative for Dengue virus RNA, and NS1.

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes and panel composition of negative Zika virus samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%#	16	20	1.Cross-reactivity panel: • Dengue PCR positive: 4 @ • Chikungunya PCR positive: 1 @ • Japanese Encephalitis positive: 1 * • West Nile Virus positive: 1 * 2.[a]Acute febrile cases: 10 3.Healthy subjects from endemic regions: 3
95%	77	80	1.Cross-reactivity panel: • Dengue PCR positive: 15 @ • Chikungunya PCR positive: 5 @ • Japanese Encephalitis positive: 5 * • West Nile Virus positive: 5 * 2.[a]Acute febrile cases: 40 3.Healthy subjects from endemic regions: 10
90%	145	150	1.Cross-reactivity panel: • Dengue PCR positive: 28 @ • Chikungunya PCR positive: 9 @ • Japanese Encephalitis positive: 9 * • West Nile Virus positive: 9 *

			2.[a]Acute febrile cases: 76 3.Healthy subjects from endemic regions: 19
85%	206	210	1.Cross-reactivity panel: • Dengue PCR positive: 40 @ • Chikungunya PCR positive: 13 @ • Japanese Encephalitis positive: 13 * • West Nile Virus positive: 13 * 2.[a]Acute febrile cases: 105 3.Healthy subjects from endemic regions: 26
80%	258	260	1.Cross-reactivity panel: • Dengue PCR positive: 49 @ • Chikungunya PCR positive: 16 @ • Japanese Encephalitis positive: 16 * • West Nile Virus positive: 16 * 2.[a]Acute febrile cases: 130 3.Healthy subjects from endemic regions: 33

[a] Acute febrile cases negative by Zika virus molecular reference assay

Positive samples / samples spiked with culture filtrate of viruses mentioned in the cross-reactivity panel may be used.

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing/dropping only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

Note:
Commercially available validated standard panels that are accepted by accreditation agencies can also be used if the clinical sample with required analyte level is not available. If PCR positive samples for cross reactive flaviviruses are not available, commercially available RNA panels/RNA from virus isolates should be used to test cross reactivity.

10. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

11. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

12. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

13. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

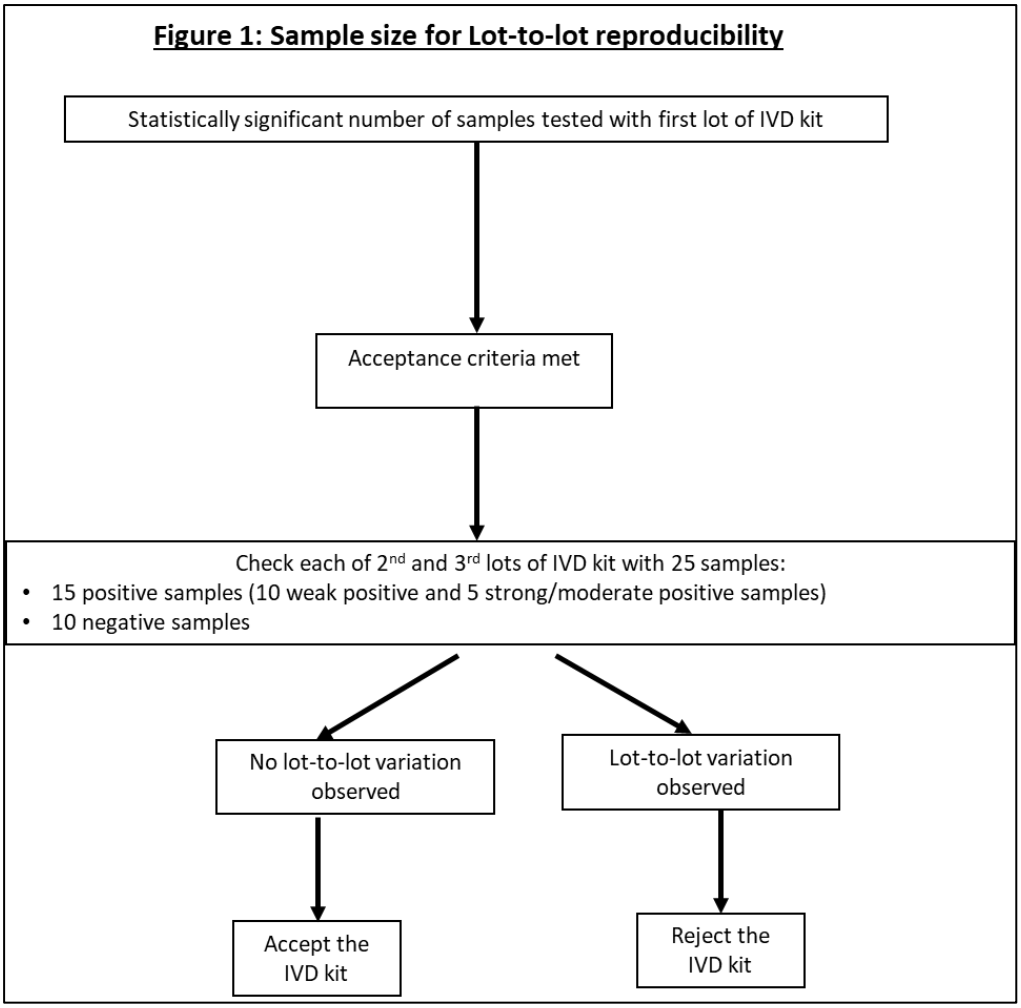
B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/ strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

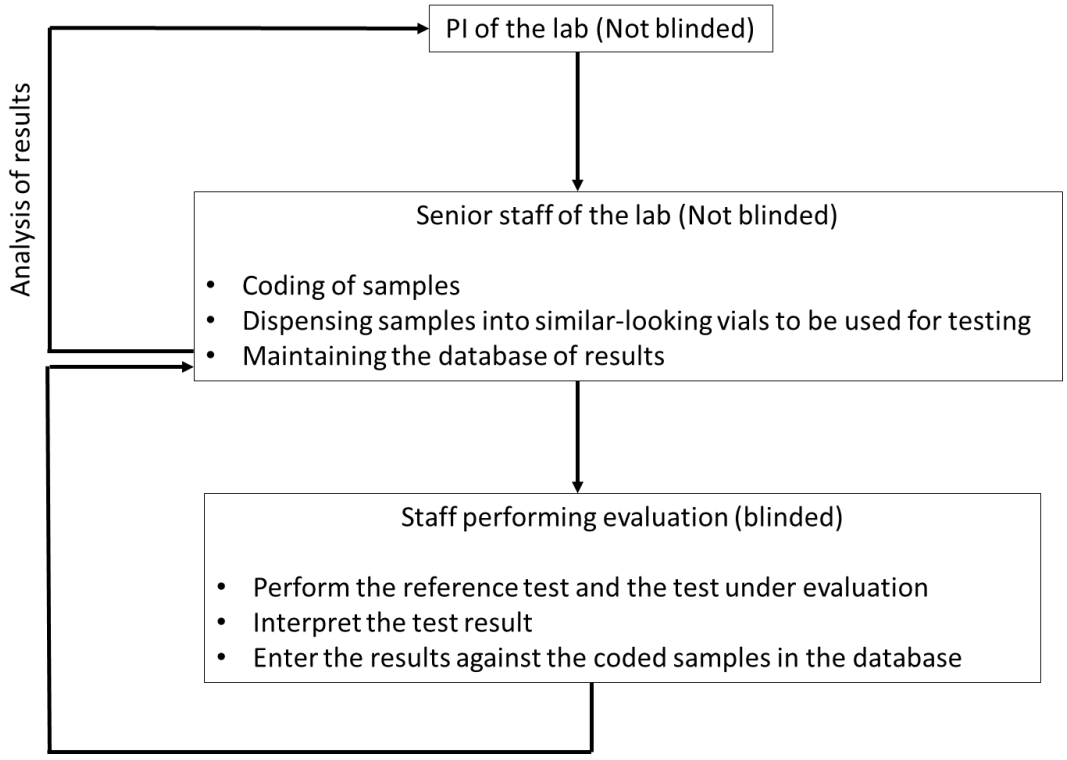
d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report Ct standard deviation across repeats (≤1cycle deviation recommended).

14. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



15. Acceptance criteria:

Sensitivity: ≥95%
Specificity: ≥98%
Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

Invalid test rate: ≤5%

Agreement between sample types– Candidate tests meant for testing multiple sample matrices should demonstrate a minimum of 95% positive percent agreement (PPA) and negative percent agreement (NPA) for all specimen types.

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥80 positive samples and ≥80 negative samples should be used for evaluation.

16. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
2. Santiago GA, Vázquez J, Courtney S, Matías KY, Andersen LE, Colón C, Butler AE, Roulo R, Bowzard J, Villanueva JM, Muñoz-Jordan JL. Performance of the Trioplex real-time RT-PCR assay for detection of Zika, Dengue, and Chikungunya viruses. Nat Commun. 2018 Apr 11;9(1):1391. doi: 10.1038/s41467-018-03772-1.
3. Stone M, Bakkour S, Grebe E, Emperador DM, Escadafal C, Deng X, Dave H, Kelly-Cirino C, Lackritz E, Rojas DP, Simmons G, Rabe IB, Busch MP. Standardized evaluation of Zika nucleic acid tests used in clinical settings and blood screening. PLoS Negl Trop Dis. 2023 Mar 17;17(3):e0011157.
4. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR ZIKA REAL-TIME PCR KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: clinical/spiked, strong, moderate, weak/simulated samples)	
	Negative samples (provide details clinical/spiked, including cross reactivity panel/simulated samples)	

Results

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Zika virus real-time PCR kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:

FINAL CONCLUSION

Performance: Satisfactory / Not satisfactory

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ..... sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for …….. Kit (Lot No……) manufactured by …… (supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal…….

DENGUE IgG BASED ASSAYS

Performance evaluation protocol for Dengue IgG RDT kits

I. Background:

CDSCO/ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding in-vitro diagnostic kit (IVD) performance.

II. Purpose:

To evaluate the performance characteristics of Dengue IgG RDT kits in the diagnosis of primary and secondary dengue infections using irreversibly de-identified leftover archived clinical samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified archived/ spiked leftover clinical samples
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through ALL of the following:
   A. Accreditation for at least one of the Quality management systems (accreditation for Testing Lab / Calibration Lab (ISO/IEC: 17025), Medical Lab (ISO: 15189), PT provider (ISO/IEC: 17043) or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel

- Handling of Dengue IgG Rapid IVD kits received for performance evaluation (Verification/Storage/Unpacking etc).
- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Dengue IgG Rapid IVD kit evaluation panel:

Well characterised Dengue IVD kit evaluation panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of sera samples should be collected from Dengue NS1/PCR/IgM confirmed cases. Further characterised for Dengue IgG positivity by using approved reference kits having high sensitivity and specificity.

Dengue IgG performance evaluation panel need to be tested again by the reference assays at the time of evaluating a particular index test to confirm the positive and negative status of the samples.

4. Reference assay:

Positive and negative samples should be characterized using composite reference standard of Dengue IgG AND one additional marker of Dengue (NS1 or IgM or PCR). The following kits should be used for characterization of the sample panel:

- *Panbio Dengue IgG capture ELISA kit/ WHO Pre-Qualified/ US-FDA/ ATAGI Australia/ PMDA Japan approved Dengue IgG ELISA kit*
- *WHO Pre-Qualified/ US-FDA/ ATAGI Australia/ PMDA Japan approved Dengue IgM ELISA kit*
- *NS1 antigen status to be assessed using WHO Pre-Qualified/ US-FDA/ ATAGI Australia/ PMDA Japan approved NS1 ELISA kit*
- *Serotype status to be assessed using a combination of CDC/NIV real-time PCR serotyping protocols.*

5. Sample size for performance evaluation:

Sample sizes of positive and negative samples and sample panel composition against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5% using the following formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Positive samples: The panel of positive samples should include samples positive for IgG by the reference assay. The samples should also be positive for either dengue NS1 antigen or dengue IgM antibodies.

It is recommended (but not mandatory) to have IgG positive samples from primary and secondary dengue cases.

Negative samples: Samples which are negative by reference dengue IgG test should form the negative sample panel.

Table 1. Sample sizes and panel composition of positive dengue IgG samples for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%	16	20	Strong Positive: 6 Moderate Positive: 7 Weak Positive: 7
95%	77	80	Strong Positive: 24 Moderate Positive: 28 Weak Positive: 28
90%	145	150	Strong Positive: 44 Moderate Positive: 53 Weak Positive: 53
85%	206	210	Strong Positive: 62

			Moderate Positive: 74 Weak Positive: 74
80%	258	260	Strong Positive: 78 Moderate Positive: 91 Weak Positive: 91
The samples need to be classified as strong, moderate and weak positives based on ELISA units of the reference assay. *#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.*			

Table 2. Sample sizes and panel composition of negative dengue IgG samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%[#]	16	20	1.Samples positive for dengue IgM/NS1/RNA but negative for IgG: **7** 2.Acute febrile illness cases: **8** • Chikungunya positive samples:2 • Dengue (NS1 & IgM & IgG & PCR) negative samples:6 3.Samples from other flavivirus disease cases (cross-reactive panel): **3** • Japanese Encephalitis IgM/IgG positive: 1 @ • West Nile Virus IgM/IgG positive: 1 * • Zika Virus IgM/IgG positive: 1 * 4. [a]Healthy subjects from endemic regions: **2**
95%	77	80	1.Samples positive for dengue IgM/NS1/RNA but negative for IgG: **27** 2.Acute febrile illness cases: **32** • Chikungunya positive samples:8 • Dengue (NS1 & IgM & IgG & PCR) negative samples:24 3.Samples from other flavivirus disease cases(cross-reactive panel): **9** • Japanese Encephalitis IgM/IgG positive: 3 @ • West Nile Virus IgM/IgG positive: 3 * • Zika Virus IgM/IgG positive: 3 * 4. [a]Healthy subjects from endemic regions: **12**
90%	145	150	1.Samples positive for dengue IgM/NS1/RNA but negative for IgG: **50** 2.Acute febrile illness cases: **60**

			• Chikungunya positive samples:15 • Dengue (NS1 & IgM & IgG & PCR) negative samples:45 3.Samples from other flavivirus disease cases(cross-reactive panel): **15** • Japanese Encephalitis IgM/IgG positive: 5 @ • West Nile Virus IgM/IgG positive: 5 * • Zika Virus IgM/IgG positive: 5 * 4. [a]Healthy subjects from endemic regions: **25**
85%	206	210	1.Samples positive for dengue IgM/NS1/RNA but negative for IgG: **70** 2.Acute febrile illness cases: **84** • Chikungunya positive samples:21 • Dengue (NS1 & IgM & IgG & PCR) negative samples:63 3.Samples from other flavivirus disease cases(cross-reactive panel): **21** • Japanese Encephalitis IgM/IgG positive: 7 @ • West Nile Virus IgM/IgG positive: 7 * • Zika Virus IgM/IgG positive: 7 * 4. [a]Healthy subjects from endemic regions: **35**
80%	258	260	1.Samples positive for dengue IgM/NS1/RNA but negative for IgG: **85** 2.Acute febrile illness cases: **104** • Chikungunya positive samples:26 • Dengue (NS1 & IgM & IgG & PCR) negative samples:78 3.Samples from other flavivirus disease cases(cross-reactive panel): **27** • Japanese Encephalitis IgM/IgG positive: 9 @ • West Nile Virus IgM/IgG positive: 9 * • Zika Virus IgM/IgG positive: 9 * 4. [a]Healthy subjects from endemic regions: **44**

[a] Samples from healthy subjects from endemic regions negative for all dengue markers (NS1, IgM, IgG, RNA)

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

Note: If IgM/IgG positive samples for cross reactive flaviviruses are not available, commercially available validated standard panels that are accepted by accreditation agencies.

6. Evaluation method:

The index test and the reference assay should be run simultaneously on the sample panel, and results should be recorded.

7. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

8. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

9. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte using the kit under evaluation.
3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

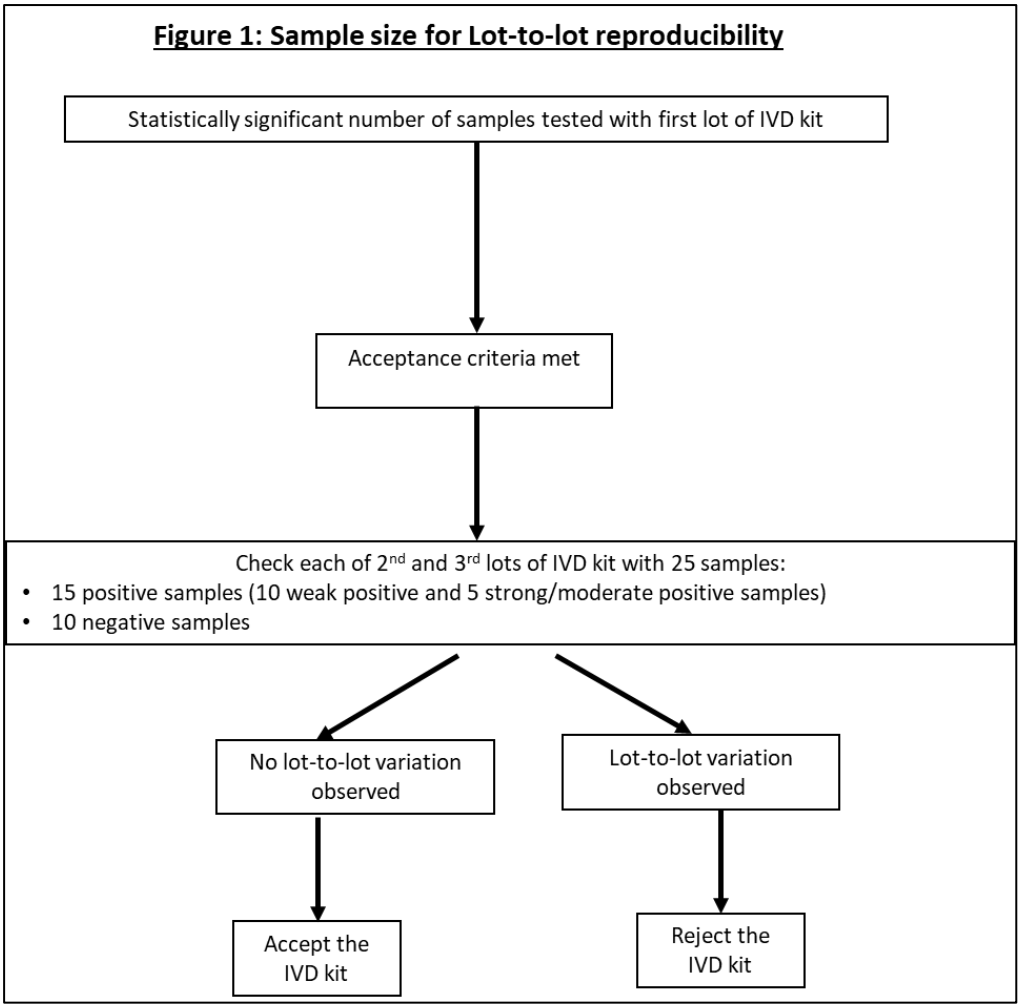
B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.



b. Inter-Operator variability: Testing and result interpretation should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

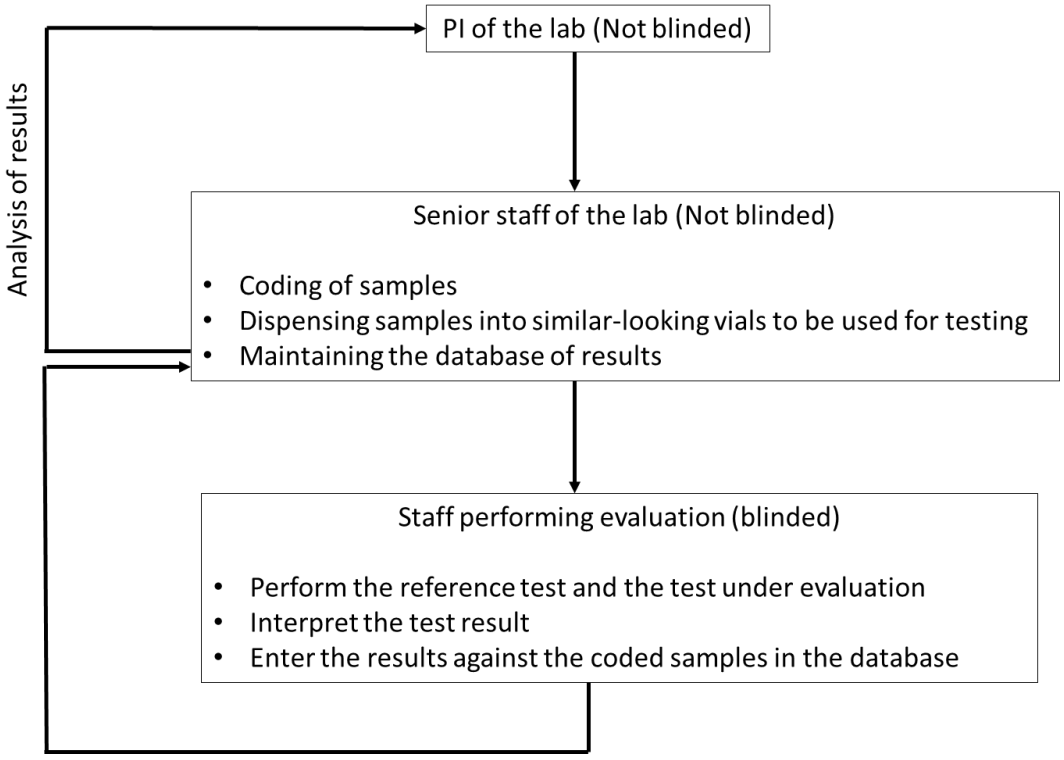
c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

10. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



11. Acceptance Criteria:

Sensitivity: ≥80%

Specificity: ≥90%

Cross reactivity: Minimal

Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥260 positive samples and ≥150 negative samples should be used for evaluation.

12. Publication Rights:

The PI (s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Vazquez S, Hafner G, Ruiz D, Calzada N, Guzman MG. Evaluation of immunoglobulin M and G capture enzyme-linked immunosorbent assay Panbio kits for diagnostic dengue infections. J Clin Virol. 2007 Jul;39(3):194-8. doi: 10.1016/j.jcv.2007.04.003..
2. WHO, Evaluation of commercially available anti-Dengue virus immunoglobulin M tests. (Diagnostics evaluation series, 3). ISBN 978 92 4 159775 3.
3. Central Drugs Standard Control Organization. Guidance on Performance Evaluation of In-vitro Diagnostic Medical Devices. 2018. Available at: https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/medical device/guidanceperformanceivd.pdf
4. Central Drugs Standard Control Organization. In-Vitro Diagnostic (IVD) Medical Devices Frequently Asked Questions. 2022. Available at: https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/IVD/FAQs/CDSCO-IVD-FAQ-03-2022-.pdf
5. U.S. Food and Drug Administration. Dengue Virus Serological Reagents - Class II Special Controls Guideline for Industry and Food and Drug Administration Staff. 2014. Available at: https://www.fda.gov/medical-devices/guidance-documents-medical-devices-and-radiation-emitting-products/Dengue-virus-serological-reagents-class-ii-special-controls-guideline-industry-and-food-and-drug
6. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
7. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

***The validation protocols need to be revisited after introduction of Dengue vaccines and the acceptance criteria needs revisiting every year so as to enable the availability of best diagnostic kits.**

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE IgG RDT KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak/simulated samples)	
	Negative samples (provide details: clinical/spiked, including cross reactivity panel/simulated samples)	

Results:

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Dengue IgG antibody - based RDT kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:

- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ….. sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for ……Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal …….

Performance evaluation protocol for Dengue IgM and IgG RDT combo kits

I. Background:

CDSCO/ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding in-vitro diagnostic kit (IVD) performance.

II. Purpose:

To evaluate the performance characteristics of Dengue IgM and IgG RDT combo kits in the diagnosis of dengue and discriminating primary and secondary dengue infections using irreversibly de-identified leftover archived clinical samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified archived/ spiked leftover clinical samples
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through ALL of the following:
   A. Accreditation for at least one of the Quality management systems (accreditation for Testing Lab / Calibration Lab (ISO/IEC: 17025), Medical Lab (ISO: 15189), PT provider (ISO/IEC: 17043) or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel

- Handling of Dengue IgM and IgG Rapid IVD kits received for performance evaluation (Verification/Storage/Unpacking etc).
- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Dengue IgM and IgG Rapid IVD kit evaluation panel:

Well characterised Dengue IVD kit evaluation panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of sera samples should be collected from Dengue NS1/PCR/IgM confirmed cases. Further characterised for Dengue IgG positivity by using approved reference kits having high sensitivity and specificity.

Dengue IgG performance evaluation panel need to be tested again by the reference assays at the time of evaluating a particular index test to confirm the positive and negative status of the samples.

4. Reference assay:

Positive and negative samples should be characterized using reference standard for Dengue IgG (and one additional marker of Dengue - NS1 or IgM or PCR) AND IgM. The following kits should be used for characterization of the sample panel:

- *Panbio Dengue IgG capture ELISA kit/ WHO Pre-Qualified/ US-FDA/ ATAGI Australia/ PMDA Japan approved Dengue IgG ELISA kit*
- *WHO Pre-Qualified/ US-FDA/ ATAGI Australia/ PMDA Japan approved Dengue IgM ELISA kit*
- *NS1 antigen status to be assessed using WHO Pre-Qualified/ US-FDA/ ATAGI Australia/ PMDA Japan approved NS1 ELISA kit*
- *Serotype status to be assessed using a combination of CDC/NIV real-time PCR serotyping protocols.*

Sample size and sample panel composition: Sample sizes of positive and negative samples of Dengue against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5% using the following formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*

- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Positive samples: The samples should be positive for dengue IgM antibodies. The panel of positive samples should include 50% of samples positive for IgG by the reference assay. Samples should be representative of varying degrees of positivity.
It is recommended (but not mandatory) to have IgG positive samples from primary and secondary dengue cases.

Negative samples: These should include samples negative by all the reference assays (True negatives).

Table 1. Sample sizes and panel composition of positive Dengue samples for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off for balanced allocation] #*	*Sample Panel Composition*
99%	16	20	1. 10 samples positive for Dengue IgM • Strong positive:3 • Moderate positive: 3 • Weak positive: 4 2. 10 samples positive for both Dengue IgM and IgG • Strong positive IgG:3 • Moderate positive IgG: 3 • Weak positive IgG: 4

95%	77	80	40 samples positive for Dengue IgM • Strong positive:12 • Moderate positive: 14 • Weak positive: 14 40 samples positive for both Dengue IgM and IgG • Strong positive IgG:12 • Moderate positive IgG: 14 • Weak positive IgG: 14
90%	145	150	75 samples positive for Dengue IgM • Strong positive:23 • Moderate positive: 26 • Weak positive: 26 75 samples positive for both Dengue IgM and IgG • Strong positive IgG: 23 • Moderate positive IgG: 26 • Weak positive IgG: 26
85%	206	210	105 samples positive for Dengue IgM • Strong positive:31 • Moderate positive: 37 • Weak positive: 37 105 samples positive for both Dengue IgM and IgG • Strong positive IgG: 31 • Moderate positive IgG: 37 • Weak positive IgG: 37
80%	258	260	130 samples positive for Dengue IgM • Strong positive:38 • Moderate positive: 46 • Weak positive: 46 130 samples positive for both Dengue IgM and IgG • Strong positive IgG: 38 • Moderate positive IgG: 46 • Weak positive IgG: 46

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes and panel composition of negative Dengue samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required [Sample size rounded off for balanced allocation] #*	*Sample Panel Composition*
99%	16	20	1.[a]Samples from acute febrile illness cases negative for dengue: **9** • Samples positive for chikungunya: 2 • Other Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR):7 2.Samples from other flavivirus disease cases (cross-reactive panel): **3** • Japanese Encephalitis IgM/IgG positive: 1@ • West Nile Virus IgM/IgG positive:1* • Zika Virus IgM/IgG positive: 1 * 3. [b]Healthy subjects from endemic regions: **8**
95%	77	80	1.[a]Samples from acute febrile illness cases negative for dengue: **44** • Samples positive for chikungunya: 8 • Other Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR):36 2.Samples from other flavivirus disease cases (cross-reactive panel): **6** • Japanese Encephalitis IgM/IgG positive: 2@ • West Nile Virus IgM/IgG positive:2* • Zika Virus IgM/IgG positive: 2 * 3. [b]Healthy subjects from endemic regions: **30**
90%	145	150	1.[a]Samples from acute febrile illness cases negative for dengue: **80** • Samples positive for chikungunya: 15

			• Other Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR):65 2.Samples from other flavivirus disease cases (cross-reactive panel): **15** • Japanese Encephalitis IgM/IgG positive: 5 @ • West Nile Virus IgM/IgG positive:5* • Zika Virus IgM/IgG positive: 5* 3. [b]Healthy subjects from endemic regions: **55**
85%	206	210	1.[a]Samples from acute febrile illness cases negative for dengue: **110** • Samples positive for chikungunya: 21 • Other Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR):89 2.Samples from other flavivirus disease cases (cross-reactive panel): **24** • Japanese Encephalitis IgM/IgG positive: 8 @ • West Nile Virus IgM/IgG positive:8* • Zika Virus IgM/IgG positive: 8* 3. [b]Healthy subjects from endemic regions: **76**
80%	258	260	1.[a]Samples from acute febrile illness cases negative for dengue: **138** • Samples positive for chikungunya: 26 • Other Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR):112 2.Samples from other flavivirus disease cases (cross-reactive panel): **27** • Japanese Encephalitis IgM/IgG positive: 9 @ • West Nile Virus IgM/IgG positive:9* • Zika Virus IgM/IgG positive: 9* 3. [b]Healthy subjects from endemic regions: **95**

[a] Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR)
[b] Samples from healthy subjects from endemic regions negative for all Dengue markers (NS1, IgM, IgG, RNA)

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

Note: If IgM/IgG positive samples for cross reactive flaviviruses are not available, commercially available validated standard panels that are accepted by accreditation agencies can also be used.

5. Evaluation method:

The index test and the reference assay should be run simultaneously on the sample panel, and results should be recorded.

6. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

7. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte using the kit under evaluation.
3 positive samples (strong, moderate and weak positive samples) and 3 negative samples for each analyte should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

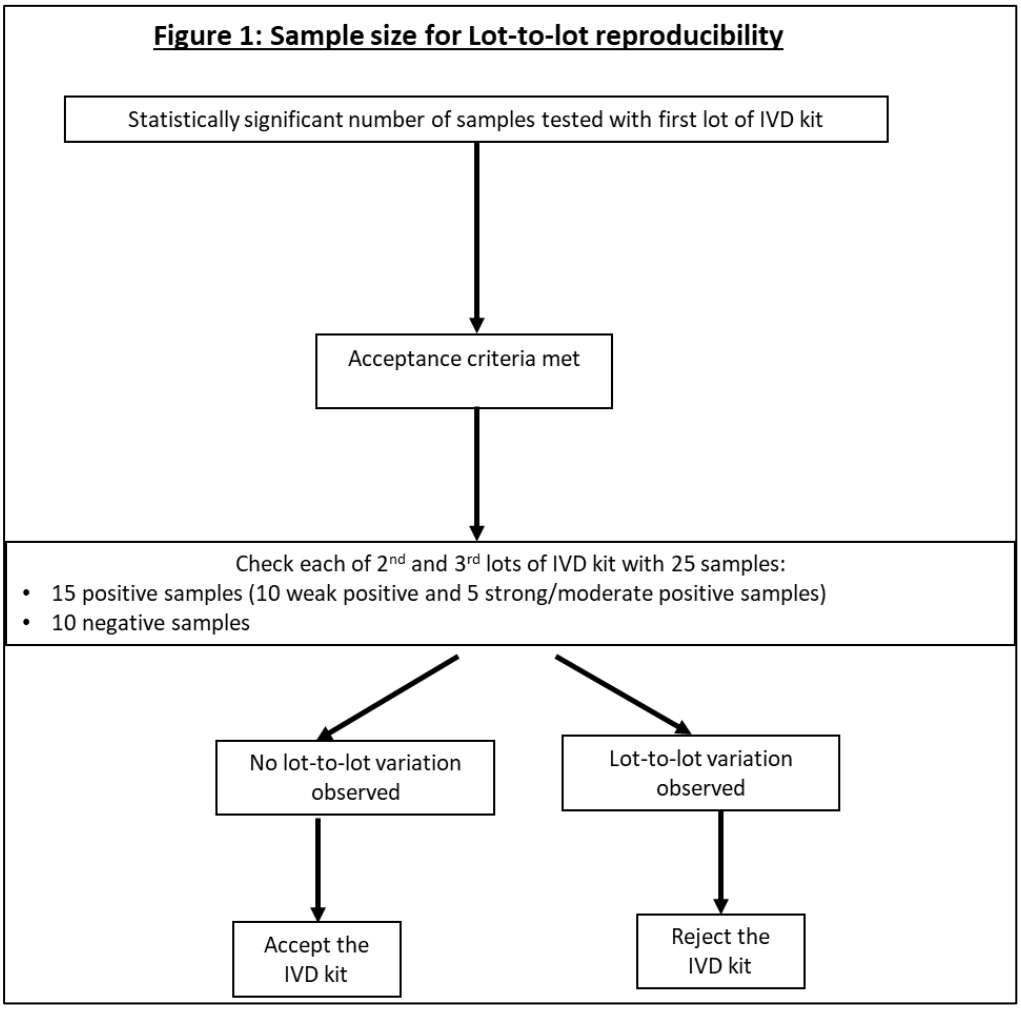
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples per target analyte, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples per target analyte, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing and result interpretation should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples for each analyte should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

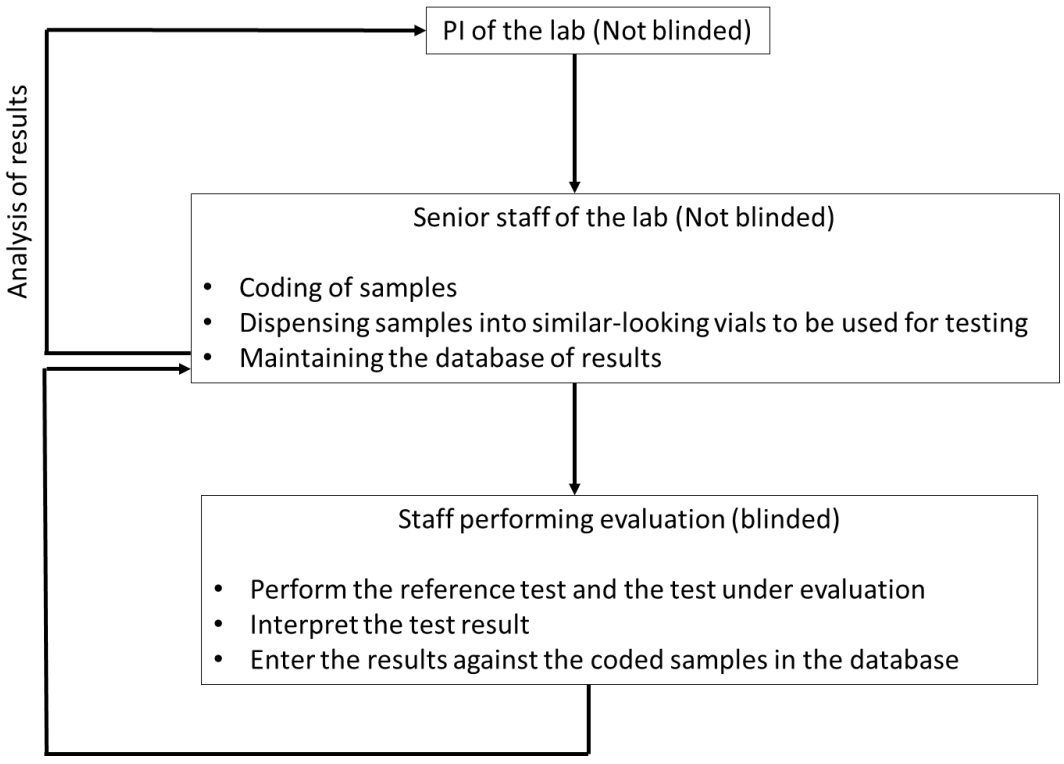
c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples for each analyte should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples for each analyte should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

8. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



9. Acceptance Criteria:

Sensitivity for each analyte: ≥80%

Specificity for each analyte: ≥90%

Cross-reactivity: Minimal

Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥260 positive samples and ≥150 negative samples for each analyte should be used for evaluation.

10. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Vazquez S, Hafner G, Ruiz D, Calzada N, Guzman MG. Evaluation of immunoglobulin M and G capture enzyme-linked immunosorbent assay Panbio kits for diagnostic dengue infections. J Clin Virol. 2007 Jul;39(3):194-8. doi: 10.1016/j.jcv.2007.04.003..
2. WHO, Evaluation of commercially available anti-Dengue virus immunoglobulin M tests. (Diagnostics evaluation series, 3). ISBN 978 92 4 159775 3.
3. Central Drugs Standard Control Organization. Guidance on Performance Evaluation of In-vitro Diagnostic Medical Devices. 2018. Available at: https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/medical device/guidanceperformanceivd.pdf
4. Central Drugs Standard Control Organization. In-Vitro Diagnostic (IVD) Medical Devices Frequently Asked Questions. 2022. Available at: https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/IVD/FAQs/CDSCO-IVD-FAQ-03-2022-.pdf
5. U.S. Food and Drug Administration. Dengue Virus Serological Reagents - Class II Special Controls Guideline for Industry and Food and Drug Administration Staff. 2014. Available at: https://www.fda.gov/medical-devices/guidance-documents-medical-devices-and-radiation-

emitting-products/Dengue-virus-serological-reagents-class-ii-special-controls-guideline-industry-and-food-and-drug
6. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
7. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

***The validation protocols need to be revisited after introduction of Dengue vaccines and the acceptance criteria needs revisiting every year so as to enable the availability of best diagnostic kits.**

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE IgM and IgG COMBO RDT KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak/simulated samples)	
	Negative samples (provide details: clinical/spiked, including cross reactivity panel/simulated samples)	

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Dengue antibody combo RDT kit	Positive			
	Negative			
	Total			

Prepare the above 2x2 table for each analyte and for overall performance characteristics

- Details of cross reactivity:
- Invalid test rate:

- **Conclusions:**
  - Sensitivity, specificity for dengue IgM:
  - Sensitivity, specificity for dengue IgG:
  - Invalid test rate:
  - Performance:
    - **Satisfactory / Not satisfactory for Dengue IgM**
    - **Satisfactory / Not satisfactory for Dengue IgG**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ..... sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for …….. Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……

Performance evaluation protocol for Dengue IgG ELISA kits

I. Background:

CDSCO/ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding in-vitro diagnostic kit (IVD) performance.

II. Purpose:

To evaluate the performance characteristics of Dengue IgG ELISA kits in the diagnosis of Dengue infection using irreversibly de-identified leftover archived clinical samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approval:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified archived/ spiked leftover clinical samples
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through ALL of the following:
   A. Accreditation for at least one of the Quality management systems (accreditation for Testing Lab / Calibration Lab (ISO/IEC: 17025), Medical Lab (ISO: 15189), PT provider (ISO/IEC: 17043) or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel

- Handling of Dengue IgG ELISA IVD kits received for performance evaluation (Verification/Storage/Unpacking etc).
- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Dengue IgG ELISA IVD kit evaluation panel:

Well characterised Dengue IVD kit evaluation panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of sera samples should be collected from Dengue NS1/PCR/IgG confirmed cases. Further characterised for Dengue IgM positivity by using approved reference kits having high sensitivity and specificity.

Dengue IgG performance evaluation panel need to be tested again by the reference assays at the time of evaluating a particular index test to confirm the positive and negative status of the samples.

4. Reference assay:

Positive and negative samples should be characterized using composite reference standard of Dengue IgG AND one additional marker of Dengue (NS1 or IgM or PCR). The following kits should be used for characterization of the sample panel:

- *Panbio Dengue IgG capture ELISA kit/ WHO Pre-Qualified/ US-FDA/ ATAGI Australia/ PMDA Japan approved Dengue IgG ELISA kit*
- *WHO Pre-Qualified/ US-FDA/ ATAGI Australia/ PMDA Japan approved Dengue IgM ELISA kit*
- *NS1 antigen status to be assessed using WHO Pre-Qualified/ US-FDA/ ATAGI Australia/ PMDA Japan approved NS1 ELISA kit*
- *Serotype status to be assessed using a combination of CDC/NIV real-time PCR serotyping protocols.*

5. Sample size for performance evaluation:

Sample sizes of positive and negative samples and sample panel composition against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, and an absolute precision of 5% using the following formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\,(1 - S_e)}{d^2}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\,(1 - S_p)}{d^2}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*

- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*

Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Positive samples: The panel of positive samples should include samples positive for IgG by the reference assay. The samples should also be positive for either dengue NS1 antigen or dengue IgM antibodies. Samples should be representative of varying degrees of positivity.
It is recommended (but not mandatory) to have IgG positive samples from primary and secondary dengue cases.

Negative samples: These should include samples negative by the reference assays for dengue IgG.

Table 1. Sample sizes and panel composition of positive Dengue samples for different values of sensitivity claimed by the manufacturer.

Sensitivity	*Calculated sample size*	*Minimum no. of Positive Samples required [Sample size rounded off] #*	*Sample Panel Composition*
99%	15	20	Strong Positive: 6 Moderate Positive: 7 Weak Positive: 7
95%	73	80	Strong Positive: 24 Moderate Positive: 28 Weak Positive: 28
90%	138	140	Strong Positive: 42 Moderate Positive: 49 Weak Positive: 49
85%	196	200	Strong Positive: 60 Moderate Positive: 70 Weak Positive: 70
80%	246	250	Strong Positive: 75 Moderate Positive: 87 Weak Positive: 88

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes and panel composition of negative Dengue samples for different values of specificity claimed by the manufacturer.

Specificity	*Calculated sample size*	*Minimum no. of Negative Samples required [Sample size rounded off]#*	*Sample Panel Composition*
99%	15	20	1.Samples positive for dengue IgM/NS1/RNA but negative for IgG: **7** 2.Acute febrile illness cases: **8** • Chikungunya positive samples:2 • Dengue (NS1 & IgM & IgG & PCR) negative samples:6 3.Samples from other flavivirus disease cases (cross-reactive panel): **3** • Japanese Encephalitis IgM/IgG positive: 1 @ • West Nile Virus IgM/IgG positive: 1 * • Zika Virus IgM/IgG positive: 1 * 4. [a]Healthy subjects from endemic regions: **2**
95%	73	80	1.Samples positive for dengue IgM/NS1/RNA but negative for IgG: **27** 2.Acute febrile illness cases: **32** • Chikungunya positive samples:8 • Dengue (NS1 & IgM & IgG & PCR) negative samples:24 3.Samples from other flavivirus disease cases(cross-reactive panel): **9** • Japanese Encephalitis IgM/IgG positive: 3 @ • West Nile Virus IgM/IgG positive: 3 * • Zika Virus IgM/IgG positive: 3 * 4. [a]Healthy subjects from endemic regions: **12**
90%	138	140	1.Samples positive for dengue IgM/NS1/RNA but negative for IgG: **45** 2.Acute febrile illness cases: **60** • Chikungunya positive samples:15 • Dengue (NS1 & IgM & IgG & PCR) negative samples:45

			3.Samples from other flavivirus disease cases(cross-reactive panel): **15** • Japanese Encephalitis IgM/IgG positive: 5 @ • West Nile Virus IgM/IgG positive: 5 * • Zika Virus IgM/IgG positive: 5 * 4. [a]Healthy subjects from endemic regions: **20**
85%	196	200	1.Samples positive for dengue IgM/NS1/RNA but negative for IgG: **65** 2.Acute febrile illness cases: **84** • Chikungunya positive samples:21 • Dengue (NS1 & IgM & IgG & PCR) negative samples:63 3.Samples from other flavivirus disease cases(cross-reactive panel): **21** • Japanese Encephalitis IgM/IgG positive: 7 @ • West Nile Virus IgM/IgG positive: 7 * • Zika Virus IgM/IgG positive: 7 * 4. [a]Healthy subjects from endemic regions: **30**
80%	246	250	1.Samples positive for dengue IgM/NS1/RNA but negative for IgG: **80** 2.Acute febrile illness cases: **104** • Chikungunya positive samples:26 • Dengue (NS1 & IgM & IgG & PCR) negative samples:78 3.Samples from other flavivirus disease cases(cross-reactive panel): **27** • Japanese Encephalitis IgM/IgG positive: 9 @ • West Nile Virus IgM/IgG positive: 9 * • Zika Virus IgM/IgG positive: 9 * 4. [a]Healthy subjects from endemic regions: **39**

[a] Acute febrile cases negative for Dengue (NS1 & IgM & IgG & PCR)
[b] Samples from healthy subjects from endemic regions negative for all Dengue markers (NS1, IgM, IgG, RNA)

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Cross reactivity panel is arranged in descending order of priority.
The pathogens marked @ are essentially to be tested.
*It is recommended to test for all pathogens listed in the cross reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

Note: If IgM/IgG positive samples for cross reactive flaviviruses are not available, commercially available validated standard panels that are accepted by accreditation agencies can also be used.

6. Evaluation method:

The index test and the reference assay should be run simultaneously on the sample panel, and results should be recorded.

7. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

8. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

9. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte(s) using the kit under evaluation.
3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

Reproducibility testing should include the following:

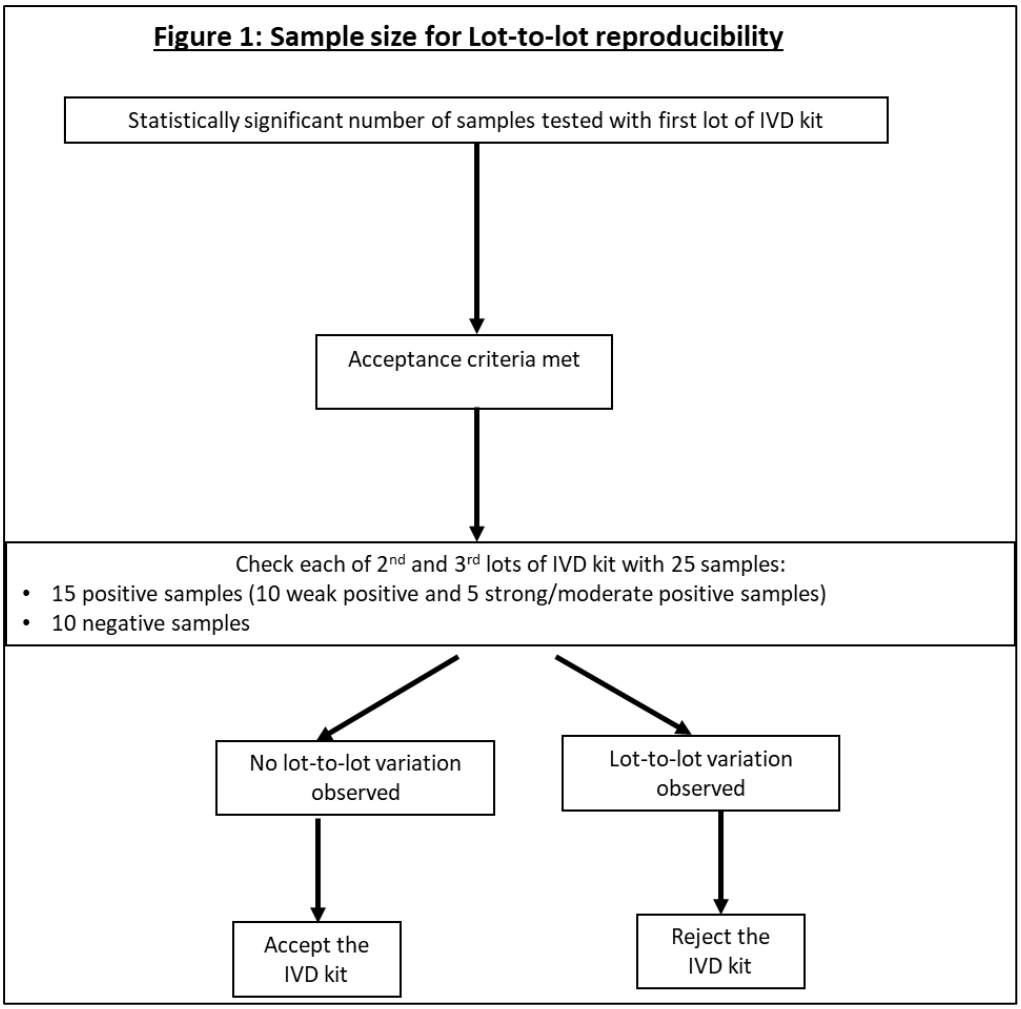
a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).

- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

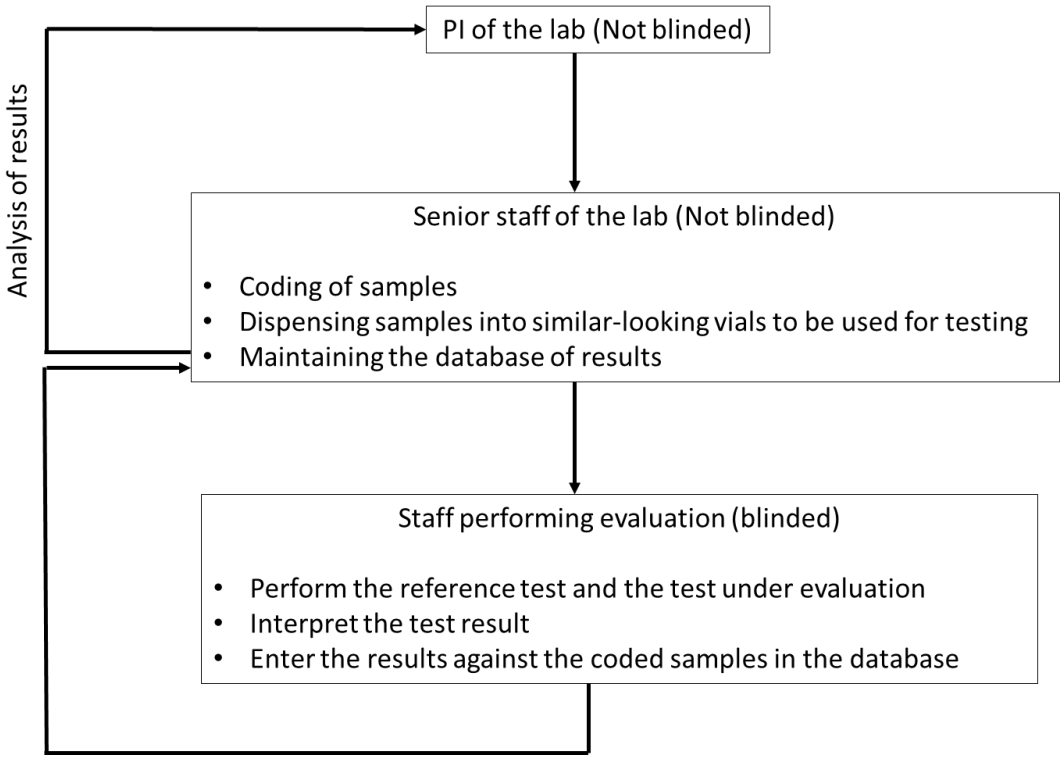
d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable).

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

10. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



11. Acceptance criteria:

Sensitivity: ≥90%

Specificity: ≥95%

Cross-reactivity: Minimal

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥140 positive samples and ≥80 negative samples should be used for evaluation.

12. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Vazquez S, Hafner G, Ruiz D, Calzada N, Guzman MG. Evaluation of immunoglobulin M and G capture enzyme-linked immunosorbent assay Panbio kits for diagnostic dengue infections. J Clin Virol. 2007 Jul;39(3):194-8. doi: 10.1016/j.jcv.2007.04.003..
2. **WHO, Evaluation of commercially available anti-Dengue virus immunoglobulin M tests. (Diagnostics evaluation series, 3). ISBN 978 92 4 159775 3.**
3. Central Drugs Standard Control Organization. Guidance on Performance Evaluation of In-vitro Diagnostic Medical Devices. 2018. Available at: https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/medical device/guidanceperformanceivd.pdf
4. Central Drugs Standard Control Organization. In-Vitro Diagnostic (IVD) Medical Devices Frequently Asked Questions. 2022. Available at: https://cdsco.gov.in/opencms/export/sites/CDSCO_WEB/Pdf-documents/IVD/FAQs/CDSCO-IVD-FAQ-03-2022-.pdf
5. U.S. Food and Drug Administration. Dengue Virus Serological Reagents - Class II Special Controls Guideline for Industry and Food and Drug Administration Staff. 2014. Available at: https://www.fda.gov/medical-devices/guidance-documents-medical-devices-and-radiation-emitting-products/Dengue-virus-serological-reagents-class-ii-special-controls-guideline-industry-and-food-and-drug
6. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
7. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

***The validation protocols need to be revisited after introduction of Dengue vaccines and the acceptance criteria needs revisiting every year so as to enable the availability of best diagnostic kits.**

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR DENGUE IgG ELISA KIT

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak/simulated samples)	
	Negative samples (provide details: clinical/spiked, including cross reactivity panel/simulated samples)	

Results:

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Dengue IgG antibody -based ELISA kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:
- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ….. sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design

2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for ……Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal……

Information on Operational and Test Performance Characteristics Required from Manufacturers for Dengue/Chikungunya/ Zika IVD

The manufacturer should provide the following details about the IVD:

1. Instructions for Use

2. Scope of the IVD: to diagnose Dengue and/or/Chikungunya and/or Zika virus

3. Intended Use Statement

4. Principle of the assay

5. Intended testing population(cases of acute febrile illness/suspected cases of Dengue and/or Chikungunya and/or Zika virus infection)

6. Intended user(laboratory professional and/or health care worker at point-of-care)

7. Detailed test protocol

8. Lot/batch No.

9. Date of manufacture

10. Date of Expiry

11. Information on operational Characteristics

    i. Configuration of the kit/device

    ii. Requirement of any additional equipment, device

    iii. Requirement of any additional reagents

    iv. Operation conditions

    v. Storage and stability before and after opening

    vi. Internal control provided or not

    vii. Quality control and batch testing data

    viii. Biosafety aspects- waste disposal requirements

10. Information on Test Performance Characteristics

    i. Type of sample-serum/plasma/whole blood/other specimen (specify)

    ii. Volume of sample

    iii. Sample rejection criteria (if any)

    iv. Any additional sample processing required

    v. Any additional device/consumable like sample transfer device, pipette, tube, etc required

    vi. Name of analyte to be detected

    vii. Pathogens targeted by the kit

viii. Time taken for testing

ix. Time for result reading and interpretation

x. Manual or automated(equipment)reading

xi. Limit of detection/ Limit of Quantification and rage of detection

xii. Diagnostic sensitivity

xiii. Diagnostic specificity

xiv. Stability and reproducibility (including data)

xv. Training required for testing (if any)

xvi. If yes, duration

xvii. Details of Cut-off and /or Equivocal Zone for interpretation of test

xviii. Details of cross reactivity, if any

xix. Interpretation of invalid and indeterminate results to be provided

xx. It is recommended to provide data demonstrating accuracy and precision

*****Please mention "Not applicable" against sections not pertaining to the kit.

INFLUENZA VIRUS, SARS-COV-2, RESPIRATORY SYNCYTIAL VIRUS IN-VITRO DIAGNOSTICS

List of Contributors:

A. Working Group:

1. Dr. Varsha Potdar, Scientist-E, ICMR-National Institute of Virology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
2. Dr. Biswajyoti Borkakoty, Scientist-F, ICMR-Regional Medical Research Center, Dibrugarh, Department of Health Research, Ministry of Health and Family Welfare, Government of India
3. Dr. Shantala GB, Professor of Microbiology, Bangalore Medical College and Research Institute, Bengaluru
4. Dr. Rizwan SA, Scientist-D, ICMR-National Institute of Epidemiology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
5. Dr. Labanya Mukhopadhyay, Scientist-C, Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

B. Review Committee:

1. Dr. Mandeep Chadha, Former Scientist-G, ICMR-National Institute of Virology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
2. Dr. Amita Jain, Former Professor and Head of the Department of Microbiology, King George's Medical University, Lucknow, Uttar Pradesh
3. Dr. Lalit Dar, Professor and Head of the Department of Microbiology, All India Institute of Medical Sciences, Delhi
4. Dr. Manoj Murhekar, Scientist-G and Director, ICMR-National Institute of Epidemiology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
5. Mr. Pramod Meshram, Deputy Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
6. Dr. Sella Senthil, Assistant Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
7. Dr. Md Omair Anwar, Drugs Inspector (Medical Devices), IVD Division, Central Drugs Standard Control Organization, Government of India
8. Dr. Nivedita Gupta, Scientist-G and Head of the Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

General Guidelines

Protocols for performance evaluation of in vitro molecular diagnostic kits for detection and differentiation of Influenza virus and/or SARS-CoV-2 and/or RSV

1. Introduction:

This document provides a framework for evaluating the performance characteristics of *in vitro* diagnostic (IVD) kits used in identifying and distinguishing various strains of Influenza viruses and/or SARS-CoV-2 and/or RSV, aligning with international standards to ensure reliability and accuracy in diagnosis. The coronavirus diseases 2019 (COVID-19) pandemic, caused by the SARS-CoV-2 virus, has necessitated the rapid development and validation of in vitro molecular diagnostic kits. These kits are crucial for the timely detection and differentiation of major respiratory viruses (influenza virus and/or its strains/SARS-CoV-2/RSV) to control their spread. This protocol outlines a systematic approach for validating these diagnostic kits to ensure their accuracy, sensitivity, specificity, and reliability.

Although SARS-CoV -2 is no longer a public health emergency globally, it is prudent to implement integrated surveillance for Influenza, SARS-CoV-2 and other respiratory viruses, making differential diagnosis for these viruses essential. Additionally, timely diagnosis of other respiratory viruses, particularly Respiratory syncytial virus (RSV), is crucial for providing effective clinical management to pediatric cases.

This document provides guidance for single plex or multiplex assays for the differential diagnosis of Influenza and/or its strains) and/or SARS-CoV-2 and/or RSV. It outlines the evaluation of IVD kits intended for the detection and differentiation of influenza virus strains and/or detection of SARS-CoV-2 and/or detection and differentiation of RSV using nucleic acid detection methods as outlined in the scope below. This includes IVD kits that detect and differentiate between influenza virus types (Influenza A or B), subtypes (A (H1N1) pdm09 or A (H3N2)), and/or multiple influenza virus types/subtypes; kits that identify only SARS-CoV-2, as well as kits that only detect and/or differentiate RSV. Additionally, this protocol may be used for multiplex IVD kits designed to simultaneously detect Influenza A & B (with or without subtyping), and/or SARS-CoV-2, and/or RSV. This document outlines the following aspects of performance evaluation of IVD kits as per the scope outlined in the document:

1.1 The procedure for validating entities to determine operational parameters of IVD kits that detect influenza virus gene segment(s).

1.2 T procedure for validating entities to determine operational parameters of IVD kits that detect SARS-CoV-2 gene segment(s).

1.3 The procedure for validating entities to determine operational parameters of IVD kits that detect RSV gene segment(s).

1.4 The techniques for identifying influenza virus/SARS-CoV-2/RSV nucleic acid targets in single-plex or multiplex formats (using appropriate protocols listed in the document).

1.5 This document does not cover performance evaluation of serological assays for detection of antigen and antibody for influenza viruses/SARS-CoV-2/RSV. The IVD kit to be validated is henceforth known as the "Kit under Evaluation."

2. Objective:

This document aims to offer a comprehensive set of instructions for evaluating the performance of molecular IVD assays mentioned in the scope below for detecting Influenza A and Influenza B viruses with/without subtyping, and other common respiratory viruses such as SARS-CoV-2 and RSV. This evaluation will focus on measuring the analytical sensitivity and specificity, cross-reactivity, repeatability, and reproducibility as compared against a reference assay using clinical sample panel.

In brief, the objectives are as follows:

2.1 To validate the performance characteristics of in vitro molecular diagnostic kits for detecting Influenza A & B (with/without subtyping)/ SARS-CoV-2/ RSV.

2.2 To ensure the kits under evaluation meet the necessary standards for sensitivity, specificity, repeatability, and reproducibility.

2.3 To evaluate the cross-reactivity of the kits with other respiratory viruses.

3. Scope:

This guideline is solely for the evaluation and establishment of the performance characteristics of IVD kits designed for the detection and subtyping of commonly circulating seasonal Influenza viruses (Influenza A(H1N1) pdm09, Influenza A(H3N2), Influenza B(Yamagata) and Influenza B(Victoria) subtypes) and/or other common respiratory viruses such as SARS-CoV-2 and RSV, using single or multiplex molecular assays (as outlined in the scope below) intended for human clinical samples. This document is a guide to assess:

3.1 The analytical assay performance characteristics with clinical specimens for the detection and/or differentiation of influenza viruses. (Protocol A)

3.2 The analytical assay performance characteristics with clinical specimens for the detection of SARS-CoV-2 (Protocol B)

3.3 The analytical assay performance characteristics with clinical specimens for the detection of RSV (Protocol C)

3.4 The analytical performance characteristics of multiplex assay for detection of two or more of these viruses by combining Protocols A, B & C as per the kit format.

3.5 Analytical performance characteristics which should include sensitivity, specificity, cross-reactivity, and lot-to-lot variation including functionality of devices that identify and/or differentiate influenza viruses, SARS-CoV-2 and/or RSV depending on the kit format.

3.6 The performance of the kit, only if the kit includes an internal control (**preferably** endogenous, or exogenous).

3.7 This document may also apply to forthcoming influenza, SARS-CoV-2 and RSV molecular diagnostic kits that do not fit within these current classifications.

3.8 The document will serve as a reference for assessing kits based on Nucleic Acid Amplification Test (single plex or multiplex assays) as listed below:

3.8.1 Real-time Reverse Transcription Polymerase Chain Reaction format (rRT-PCR): including Real-time PCR probe-based assays or non-probe based assays

3.8.2 Other NAT testing platforms such as LAMP/RPA, and other closed system platforms such as TrueNat /cartridge-based assays

Note: *This protocol is not suitable for the kits where amplicons are handled outside the amplification system.*

4. Requirements:

4.1 Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment and consumables.

4.2 Evaluation sites/laboratories (With required equipment)

4.4 Reference test kits

4.4 Characterized clinical samples for evaluation panel

4.5 Laboratory supplies

5. Ethical approvals:

Laboratory validation of IVDs using irreversibly de-identified samples is exempted from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024. A self-declaration form as provided in ICMR guidelines to be submitted by the investigators to the institutional authorities and ethics committee for information (https://ethics.ncdirindia.org/asset/pdf/Guidance_on_Ethical_Requirements_for_Laboratory_Validation_Testing.pdf)

6. Procedure:

6.1 Study design/type: Diagnostic accuracy study using leftover irreversibly de-identified archived clinical samples.

6.2 Evaluation site/laboratory considerations: Identified IVD kit evaluation laboratories should establish their proficiency through

6.2.1 Accreditation for at least one of the Quality management systems (accreditation for Testing Lab / Calibration Lab (ISO/IEC: 17025), Medical Lab (ISO 15189), PT provider (ISO/IEC: 17043) or CDSCO approved Reference laboratory.

6.2.2 Have sufficient numbers of archived as well as contemporary (collected and tested within the preceding 1 year) clinical specimens positive for respiratory viruses targeted by the kit under evaluation (Influenza A(H1N1)pdm09, A(H3N2), B(Yamagata), B(Victoria), and/or SARS-CoV-2 and/or RSV A & B), with aliquots stored at -80 ℃ deep freezers or in lyophilized form.

6.2.3 Virus strains should be well-characterized by ICMR approved or US FDA/ ATAGI Australia/PMDA Japan approved/WHO Pre-Qualified reference assay and/or by influenza virus HA gene/segment or gene-specific sequencing (for SARS-CoV-2 and RSV) or Next-Generation Sequencing.

6.2.4 Have a minimum BSL-2 level facility with trained manpower and at least two different Real Time platforms to perform molecular diagnostic assays for Influenza virus and other respiratory viruses.

6.2.5 Have a good record of External Quality Assurance programs for influenza, SARS-CoV-2, and other respiratory viruses.

6.2.6 Staff training: All the staff involved in IVD kit evaluation should undergo hands-on training and competency testing on the following:

6.2.6.1 Preparation & characterization of kit evaluation panel

6.2.6.2 Handling of respiratory virus PCR kits received for performance evaluation (Verification/Storage/Unpacking etc).

6.2.6.3 Testing, interpreting, recording of results & reporting

6.2.6.4 Data handling, data safety & confidentiality

6.3 Performance characteristics: To be assessed for all assay targets of influenza A/B, SARS-CoV-2 and RSV (single plex or multi-plex assays)

6.3.1 Analytical Sensitivity and specificity
6.3.2 Cross-reactivity
6.3.3 Repeatability
6.3.4 Reproducibility

Protocol A

Performance Evaluation of Molecular IVD Kit detecting influenza A & B viruses, and subtyping into A (H1N1) pdm 09, A(H3N2), B(Yamagata) & B(Victoria) in single plex or multiplex format

1. **Objective:**

   1.1 To evaluate the performance of molecular IVD kit for detection and differentiation of Influenza viruses as per the scope outlined in this document.

   1.2 To ensure the kits under evaluation meet the necessary standards for sensitivity, specificity, repeatability, and reproducibility.

   1.3 To evaluate the cross-reactivity of the kits with other respiratory viruses.

2. **Evaluation of performance characteristics should be done for the following parameters:**

   2.1 Sensitivity and specificity

   2.2 Cross-reactivity

   2.3 Repeatability

   2.4 Reproducibility

3. **Panel development: Clinical sample (archived/contemporary) panel for testing:**

   3.1 Contemporary (collected and tested within the preceding 1 year) leftover irreversibly de-identified clinical/archived respiratory samples (in VTM) for the panel should be irreversibly de-identified.

   3.2 Samples to be used for panel preparation shall be stored properly at – 80 ℃ or lyophilized.

   3.3 Unless the manufacturer has specific requirement of nucleic acid extraction kit, the validation laboratory can use WHO Pre-Qualified/ US FDA/ ATAGI Australia/ PMDA Japan approved/ICMR validated total RNA / viral RNA extraction kits for the evaluation.

   3.4 Clinical samples for evaluation should be characterized by a reference kit / Sequencing/NGS.

   3.5 All positive samples should be confirmed positive for the target pathogens by the reference assay.

   3.6 All negative samples should be confirmed negative for the target pathogens by the reference assay.

4. Sample size and sample panel composition for evaluation of performance characteristics:

Sample sizes of positive and negative samples of the analyte/pathogen targeted by the kit against different values of sensitivity and specificity are provided in Table 1. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity. Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ \left(1 - S_p\right)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the number of positive samples.*
- *n (sp) is the number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Sample sizes for positive samples and their composition for evaluating subtyping are provided in Table 2.

Table 1. Sample sizes per target pathogen for different values of sensitivity/ specificity claimed by the manufacturer.

Sensitivity/ Specificity	*Sample size: Minimum number of positive samples*[¥]	*Composition of positive samples*[#]	*Sample size: Minimum number of negative samples (rounded)* [¥]	*Minimum number of cross reactive* samples among the negative samples*
99%	16 (rounded to 20 for better distribution of samples)	Strong positive = 06 Moderate positive = 07 Weak positive = 07	20	5
95%	77 (rounded to 80 for better distribution of samples)	Strong positive = 24 Moderate positive = 28 Weak positive = 28	80	20
90%	146 (rounded to 155 for better distribution of samples)	Strong positive = 45 Moderate positive = 55 Weak positive =55	150	38
85%	207 (rounded to 215 for better distribution of samples)	Strong positive = 63 Moderate positive = 76 Weak positive = 76	210	53
80%	259 (rounded to 260 for better distribution of samples)	Strong positive = 78 Moderate positive = 91 Weak positive = 91	260	65
[#]Strong positive: (Ct value <25) Moderate positive: (Ct value between 25-30) Weak positive: (Ct value >30 and ≤ 34)				
[¥] Equal distribution of positive nasopharyngeal and/or oropharyngeal swabs in virus transport medium (VTM) to be used				
* Samples positive for common respiratory viruses (such as SARS-CoV-2, Parainfluenza viruses, Adenoviruses, Rhinoviruses, Respiratory Syncytial Virus (including its types and subtypes), common human coronaviruses), other than the ones targeted by the kit under evaluation. Equal distribution of cross-reactive viruses is desirable. Commercially available validated standard panels that are accepted by accreditation agencies can also be used if clinical samples with required target is not available for cross-reactivity analysis.				

It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Sample sizes for positive samples and their composition for evaluating subtyping

	Sample size* (per target pathogen)			
	Influenza A (H1N1) pdm09	Influenza A/H3N2	Influenza B (Victoria)	
Sensitivity	Minimum number of nasopharyngeal swabs/ oropharyngeal swabs (rounded figures)	Minimum number of nasopharyngeal swabs/ oropharyngeal swabs (rounded figures)	Minimum number of nasopharyngeal swabs/ oropharyngeal swabs (rounded figures)	Minimum total number of positive samples (rounded figures)
99%	20	20	20	60
95%	80	80	80	240
90%	150	150	150	450

85%	210	210	210	630
80%	260	260	260	780
*Combination of strong, moderate and weak positive samples should be considered as per the information provided in Table 1.				

It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Influenza B (Yamagata) is not mentioned in the table in view nil/minimal global and national circulation of the strain for the past few years. Samples positive for influenza B (Yamagata) should be used for performance evaluation of kits if the strain starts circulating in the country.

5. Methodology:

5.1 Samples should be tested in parallel with the Kit Under Evaluation and the reference assay. The ICMR-NIV RT-qPCR assay for Influenza/SARS-CoV-2 or WHO Pre-Qualified/ US FDA/ ATAGI Australia/ PMDA Japan approved kit will be considered as the reference assay for these parameters.

5.2 The validation laboratory can use the established total RNA / viral RNA extraction protocol for the evaluation.

5.3 The instruction for the assay setup and the interpretation of the results will be as per the protocol outlined by the manufacturer of the reference test and the kit under evaluation.

5.4 The results shall be compared with the reference assay for sensitivity and specificity calculations.

5.5 If there is a discrepancy observed in the results with the index test, this discrepancy should be taken as discordant. Repetition of the assay may introduce bias. If the reference kit itself has failed, then these samples with discrepancies should be discarded, and new well-characterized samples should be used instead.

True positive samples: These are samples positive by both reference assay and index test.

True negative samples: These are samples negative by both reference assay and index test.

False positive samples: These are samples negative by reference assay and positive by index test.

False negative samples: These are samples positive by reference assay and negative by index test.

5.6 The interpretation for internal control (**preferably** endogenous, or exogenous) will be as per manufacturer's instruction.

5.7 PCR should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Real-time closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.
The details on the Real-time Equipment used for validation should be recorded, including calibration status.

6. Cross-reactivity Analysis:

6.1 Objective:

To assess the primer-probe set for true detection of influenza viruses and assess its cross-reactivity with other respiratory viruses.

6.2 Methodology:

6.2.1 Potential cross-reactivity of the kit shall be ruled out by testing other respiratory pathogen positive samples as part of the negative sample panel, which is outlined in Table 1.*

6.2.2 Cross-reactivity will be assessed by comparing the results of these samples using kit under evaluation and reference kit.

** For multiplex assays targeting influenza, SARS-CoV-2, and RSV, samples positive for these viruses may be suitably interchanged for assessing cross-reactivity, apart from the ORV panel. (i.e. Influenza A positive samples may be used for detecting cross-reactivity against Influenza B)*

7. Acceptance criteria for the kit:

Sensitivity for each pathogen/ type/ subtype: ≥95%

Specificity for each pathogen/ type/ subtype: ≥99%

Cross-reactivity: Minimal

Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥80 positive samples and ≥20 negative samples should be tested for evaluation for each pathogen/ type/ subtype.

8. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples per target).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples per target).
- There should be no lot-to-lot variability (qualitative).

b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report Ct standard deviation across repeats (≤1cycle deviation recommended).

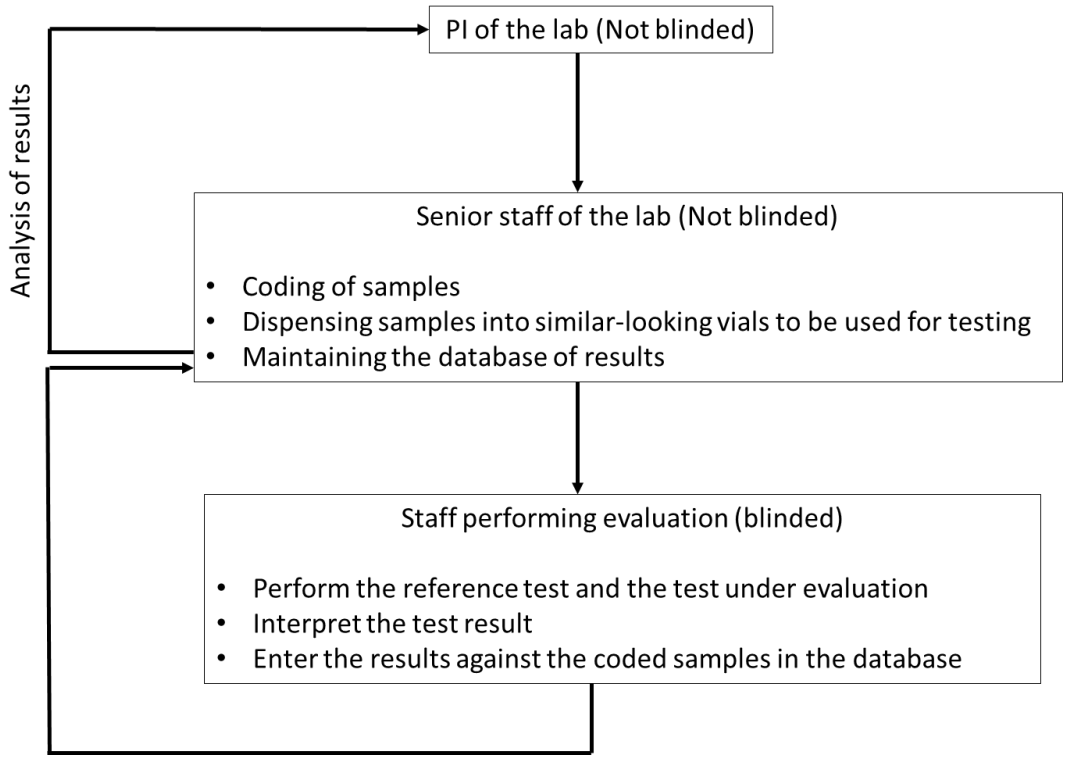
9. Internal Control Analysis:

9.1 Monitor the internal control (preferably RNaseP or other housekeeping gene) to ensure consistent extraction and amplification efficiency across samples and runs.
9.2 Ct-values of internal controls should be within the manufacturer's prescribed limit.
9.3 Tests will be marked invalid if Ct-values are outside the prescribed limit.

10. Blinding of Laboratory Staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 1.

Fig.1: Blinding in evaluation exercise



11. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

12. Conclusion:

Based on the comprehensive evaluation conducted, the [Kit & Manufacturer's Name] Influenza Virus molecular IVD kit has been found **[Satisfactory/Not Satisfactory**] for its intended *in vitro* diagnostic (IVD) use.

The assay demonstrates [Strengths/Concerns] in terms of sensitivity, specificity, and performance characteristics compared to established reference IVD approved RT-PCR kits.

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

13. Performance evaluation report format

The performance evaluation report format (page 34) is designed for multiplex assays with several targets. It should be modified and used accordingly for single plex assays/multiplex assays with fewer targets.

Protocol B:

Performance Evaluation of Molecular IVD Kit detecting SARS-CoV-2 in single plex or multiplex format

1. Objective:

1.1. To validate the performance characteristics of in vitro molecular diagnostic kits for detecting SARS-CoV-2 as per the scope outlined in this document.

1.2. To ensure the kits under evaluation meet the necessary standards for sensitivity, specificity, repeatability, and reproducibility.

1.3. To evaluate the cross-reactivity of the kits with other respiratory viruses.

2. Evaluation of Performance characteristic should be done for the following:

2.1 Sensitivity and specificity

2.2 Cross-reactivity

2.3 Repeatability

2.4 Reproducibility

3. Panel development: Clinical sample (archived/ contemporary) panel for testing:

3.1 Contemporary (collected and tested within the preceding 1 year) leftover irreversibly de-identified clinical/archived respiratory samples in VTM for the panel should be irreversibly de-identified.

3.2 Samples to be used for panel preparation shall be stored properly at – 80 ℃ or lyophilized.

3.3 Unless the manufacturer has specific requirement of nucleic acid extraction kit, the MDTLs/ validation laboratory can use WHO Pre-Qualified/ US FDA/ ATAGI Australia/ PMDA Japan approved/ICMR validated an established total RNA / viral RNA extraction kits for the evaluation.

3.4 Clinical samples for evaluation should be characterized by a reference kit / Sequencing/NGS.

3.5 All positive samples should be confirmed positive for the target pathogens by the reference assay.

3.6 All negative samples should be confirmed negative for the target pathogens by the reference assay.

4. Sample size and sample panel composition for evaluation of performance characteristics:

Sample sizes of positive and negative samples of SARS-CoV-2 against different values of sensitivity and specificity are provided in Table 3. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size that is outlined for 85% specificity. Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\,(1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\,(1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the number of positive samples.*
- *n (sp) is the number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Table 3. Sample sizes for different values of sensitivity/ specificity claimed by the manufacturer.

Sensitivity/ Specificity	*Sample size: Minimum number of positive samples*[¥]	*Composition of positive samples*[#]	*Sample size: Minimum number of negative samples (rounded)*[¥]	*Minimum number of cross reactive* samples among the negative samples*
99%	16 (rounded to 20 for better distribution of samples)	Strong positive = 06 Moderate positive = 07 Weak positive = 07	20	5

95%	77 (rounded to 80 for better distribution of samples)	Strong positive = 24 Moderate positive = 28 Weak positive = 28	80	20
90%	146 (rounded to 155 for better distribution of samples)	Strong positive = 45 Moderate positive = 55 Weak positive =55	150	38
85%	207 (rounded to 215 for better distribution of samples)	Strong positive = 63 Moderate positive = 76 Weak positive = 76	210	53
80%	259 (rounded to 260 for better distribution of samples)	Strong positive = 78 Moderate positive = 91 Weak positive = 91	260	65
[#]Strong positive: (Ct value <25) Moderate positive: (Ct value between 25-30) Weak positive: (Ct value >30 and ≤ 34)				
[¥] Nasopharyngeal/ oropharyngeal swabs in virus transport medium (VTM) to be used				
*Samples positive for common respiratory viruses (such as Influenza (including its types and subtypes), Parainfluenza viruses, Adenoviruses, Rhinoviruses, Respiratory Syncytial Virus (including its types and subtypes), common human coronaviruses), other than the ones targeted by the kit under evaluation. Equal distribution of cross-reactive viruses is desirable. Commercially available validated standard panels that are accepted by accreditation agencies can also be used if clinical samples with required target is not available for cross-reactivity analysis.				

It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

4.1 Repeatability testing will be performed on 3 positive (strong, moderate and weak positive) and 3 negative samples (within the selected positive and negative samples) per target pathogen 5 times (replicates of 5).

5. Methodology:

5.1 Samples should be tested in parallel with the Kit Under Evaluation and the reference assay. The ICMR-NIV RT-qPCR assay for Influenza/SARS-CoV-2 or WHO Pre-Qualified/ US FDA/ PMDA Japan/ ATAGI Australia approved kit will be considered as the reference assay for these parameters.

5.2 The validation laboratory can use established total RNA / viral RNA extraction protocol for the evaluation.

5.3 The instruction for the assay setup and the interpretation of the results will be as per the protocol outlined by the manufacturer of the reference test and the kit under evaluation. The results shall be compared with the reference assay for sensitivity and specificity calculations.

5.4 If there is a discrepancy observed in the results with the index test, this discrepancy should be taken as discordant. Repetition of the assay may introduce bias. If the reference kit itself has failed, then these samples with discrepancies should be discarded, and new well-characterized samples should be used instead.

True positive samples: These are samples positive by both reference assay and index test.

True negative samples: These are samples negative by both reference assay and index test.

False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

5.5 The interpretation for internal control (**preferably** endogenous, or exogenous) will be as per manufacturer's instruction.

5.6 PCR should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Real-time closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.

The details on the Real-time Equipment used for validation should be recorded, including calibration status.

The details on the Real-time Equipment used for validation should be recorded including calibration status.

6. Cross-reactivity Analysis:

6.1 Objective:

To assess the primer-probe set for true detection of SARS-CoV-2 and assess its cross-reactivity with other respiratory viruses.

6.2 Methodology:

6.2.1 Potential cross-reactivity of the kit shall be ruled out by testing other respiratory pathogen positive samples as part of the negative sample panel, as outlined in Table 3.*

6.2.2 Cross-reactivity will be assessed by comparing the results of these samples using kit under evaluation and reference kit.

** For multiplex assays targeting influenza, SARS-CoV-2, and RSV detection, samples positive for these viruses may be suitably interchanged for assessing cross-reactivity*

7. Acceptance criteria for the kit:

Sensitivity: ≥95%
Specificity: ≥99%
Cross-reactivity: Minimal
Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥80 positive samples and ≥20 negative samples should be tested for evaluation for each pathogen/ type/ subtype.

8. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples per target).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples per target).
- There should be no lot-to-lot variability (qualitative).

b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report Ct standard deviation across repeats (≤1cycle deviation recommended).

9. Internal Control Analysis:

9.1 Monitor the internal control (preferably RNaseP or other housekeeping gene) to ensure consistent extraction and amplification efficiency across samples and runs.

9.2 Ct-values of internal controls should be within the manufacturer's prescribed limit.

9.3 Tests will be marked invalid if Ct-values are outside the prescribed limit.

10. Blinding of Laboratory Staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 1 in Protocol A.

11. Conclusion:

Based on the comprehensive evaluation conducted, the [Kit & Manufacturer's Name] SARS-CoV-2 molecular IVD kit has been found **[Satisfactory/Not Satisfactory**] for its intended *in vitro* diagnostic (IVD) use.

The assay demonstrates [Strengths/Concerns] in terms of sensitivity, specificity, and performance characteristics compared to established reference IVD approved RT-PCR kits.

12. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

13. Performance evaluation report format:

The performance evaluation report format (page 34) is designed for multiplex assays with several targets. It should be modified and used accordingly for single plex assays/multiplex assays with fewer targets.

Protocol C:

Performance Evaluation of Molecular IVD Kit detecting Respiratory Syncytial Virus (RSV /RSV A/RSV B) in single plex or multiplex format

1. Objective:

1.1 To validate the performance characteristics of in vitro molecular diagnostic kits for detecting and/or differentiating RSV A/B as per the scope outlined in this document.

1.2 To ensure the kits under evaluation meet the necessary standards for sensitivity, specificity, repeatability, and reproducibility.

1.3 To evaluate the cross-reactivity of the kits with other respiratory viruses.

2. Evaluation of Performance characteristic should be done for the following:

2.1 Sensitivity and specificity

2.2 Cross-reactivity

2.3 Repeatability

2.4 Reproducibility

3. Panel development: Clinical sample (archived/ contemporary) panel for testing:

3.1 Contemporary (collected and tested within the preceding 1 year) leftover irreversibly de-identified clinical/archived respiratory samples in VTM for the panel should be irreversibly de-identified.

3.2 Samples to be used for panel preparation shall be stored properly at – 80 ℃ or lyophilized.

3.3 Unless the manufacturer has specific requirement of nucleic acid extraction kit, the MDTLs/ validation laboratory can use WHO Pre-Qualified/ US FDA/ ATAGI Australia/ PMDA Japan approved/ ICMR validated an established total RNA / viral RNA extraction kits for the evaluation.

3.4 Clinical samples for evaluation should be characterized by a reference kit / Sequencing/NGS.

3.5 All positive samples should be confirmed positive for the target pathogens by the reference assay.

3.6 All negative samples should be confirmed negative for the target pathogens by the reference assay.

4. Sample size and sample panel composition for evaluation of performance characteristics:

Sample sizes of positive and negative samples of the RSV A/B against different values of sensitivity and specificity are provided in Table 4. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%.

Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity. Sample sizes for positive samples and their composition for evaluating subtyping (RSV A/B) are provided in Table 5. Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\,(1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\,(1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the number of positive samples.*
- *n (sp) is the number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Table 4. Sample sizes per target pathogen (RSV A/B) for different values of sensitivity/ specificity claimed by the manufacturer.

Sensitivity/ Specificity	*Sample size: Minimum number of positive samples*[¥]	*Composition of positive samples*[#]	*Sample size: Minimum number of negative samples (rounded)*[¥]	*Minimum number of cross reactive* samples among the negative samples*
99%	16 (rounded to 20 for better distribution of samples)	Strong positive = 06 Moderate positive = 07 Weak positive = 07	20	5
95%	77 (rounded to 80 for better distribution of samples)	Strong positive = 24 Moderate positive = 28 Weak positive = 28	80	20

90%	146 (rounded to 155 for better distribution of samples)	Strong positive = 45 Moderate positive = 55 Weak positive =55	150	38
85%	207 (rounded to 215 for better distribution of samples)	Strong positive = 63 Moderate positive = 76 Weak positive = 76	210	53
80%	259 (rounded to 260 for better distribution of samples)	Strong positive = 78 Moderate positive = 91 Weak positive = 91	260	65
#Strong positive: (Ct value <25) Moderate positive: (Ct value between 25-30) Weak positive: (Ct value >30 and ≤ 34)				
¥ Nasopharyngeal/ oropharyngeal swabs in virus transport medium (VTM) to be used				
*Samples positive for common respiratory viruses (such as Influenza (including its types and subtypes), SARS-CoV-2, Parainfluenza viruses, Adenoviruses, Rhinoviruses, common human coronaviruses), other than the ones targeted by the kit under evaluation. Equal distribution of cross-reactive viruses is desirable. Commercially available validated standard panels that are accepted by accreditation agencies can also be used if clinical samples with required target is not available for cross-reactivity analysis.				

It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 5. Sample sizes for positive samples and their composition for evaluating subtyping

Sensitivity	Sample size* (per target pathogen)	RSV A Minimum number of nasopharyngeal swabs/ oropharyngeal swabs	RSV B Minimum number of nasopharyngeal swabs/ oropharyngeal swabs	Minimum total positive samples
99%	20	20	20	40
95%	80	80	80	160
90%	150	150	150	300
85%	210	210	210	420
80%	260	260	260	520
*Combination of strong, moderate and weak positive samples should be considered as per the information provided in Table 4.				

It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

4.1 Repeatability testing will be performed on 3 positive (strong, moderate and weak positive) and 3 negative samples (within the selected positive and negative samples) per target pathogen 5 times (replicates of 5).

5. Methodology:

5.1 Samples should be tested in parallel with the Kit Under Evaluation and the reference assay. The ICMR-NIV RT-qPCR assay for RSV or WHO Pre-Qualified/ US FDA/

ATAGI Australia/ PMDA Japan approved kit will be considered as the reference assay for these parameters.

5.2 The validation laboratory can use established total RNA / viral RNA extraction protocol for the evaluation.

5.3 The instruction for the assay setup and the interpretation of the results will be as per the protocol outlined by the manufacturer of the Kit Under Evaluation.

5.4 The results shall be compared with the reference assay for sensitivity and specificity calculations.

5.5 If there is a discrepancy observed in the results with the index test, this discrepancy should be taken as discordant. Repetition of the assay may introduce bias. If the reference kit itself has failed, then these samples with discrepancies should be discarded, and new well-characterized samples should be used instead.

True positive samples: These are samples positive by both reference assay and index test.
True negative samples: These are samples negative by both reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

5.6 The interpretation for internal control (**preferably** endogenous, or exogenous) will be as per manufacturer's instruction.

5.7 PCR should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.
Real-time closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.
The details on the Real-time Equipment used for validation should be recorded, including calibration status.
The details on the Real-time Equipment used for validation should be recorded including calibration status.

6. Cross-reactivity Analysis:

6.1 Objective:

To assess the primer-probe set for true detection of RSV and assess its cross-reactivity with other respiratory viruses.

6.2 Methodology:

6.2.1 Potential cross-reactivity of the kit shall be ruled out by testing other respiratory pathogen positive samples as part of the negative sample panel, as outlined in Table 4.*

6.2.2 Cross-reactivity will be assessed by comparing the results of these samples using kit under evaluation and reference kit.

** For multiplex assays targeting influenza, SARS-CoV-2, and RSV detection, samples positive for these viruses may be suitably interchanged for assessing cross-reactivity*

7. Acceptance criteria for the kit:

Sensitivity for each pathogen/ type/ subtype: ≥95%
Specificity for each pathogen/ type/ subtype: ≥99%
Cross-reactivity: Minimal
Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥80 positive samples and ≥20 negative samples should be tested for evaluation for each pathogen/ type/ subtype.

8. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples per target).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples per target).
- There should be no lot-to-lot variability (qualitative).

b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong,

moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative) should be 100% based on positive and negative test result (qualitative).

It is desirable to report Ct standard deviation across repeats (≤1cycle deviation recommended).

9. Internal Control Analysis:

9.1 Monitor the internal control (RNaseP or other endogenous housekeeping gene) to ensure consistent extraction and amplification efficiency across samples and runs.

9.2 Ct-values of internal controls should be within the manufacturer's prescribed limit.

9.3 Tests will be marked invalid if Ct-values are outside the prescribed limit.

10. Blinding of Laboratory Staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 1 in Protocol A.

11. Conclusion:

Based on the comprehensive evaluation conducted, the [Kit & Manufacturer's Name] RSV molecular IVD kit has been found **[Satisfactory/Not Satisfactory**] for its intended *in vitro* diagnostic (IVD) use.

The assay demonstrates [Strengths/Concerns] in terms of sensitivity, specificity, and performance characteristics compared to established reference IVD approved RT-PCR kits.

12. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

13. Performance evaluation report format:

The performance evaluation report format (page 34) is designed for multiplex assays with several targets. It should be modified and used accordingly for single plex assays/multiplex assays with fewer targets.

Protocol D:

Performance Evaluation of Molecular IVD Kit detecting Influenza virus and SARS-CoV-2 in multiplex format

To assess the performance of multiplex assays, Protocols A and B can be used as per kit format to check the performance of each virus for its sensitivity and specificity assessment, including cross reactivity, repeatability, reproducibility and Lot to lot variation.

A comprehensive report can be generated which will include sensitivity and specificity for all targets.

Sample size for multiplex molecular assay (as per the scope outlined in the document) detecting Influenza virus and SARS-CoV-2 in multiplex format is given below. All other relevant performance criteria (such as repeatability/ reproducibility/ cross-reactivity) outlined in the single plex protocols (Protocols A and B) are to be essentially followed and met for each target pathogen/type/subtype.

1. **Sample size and sample panel composition for evaluation of performance characteristics**:

Sample sizes of positive and negative samples against different values of sensitivity and specificity are provided in Table 6. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity. Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the number of positive samples.*
- *n (sp) is the number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*

- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Table 6. Sample sizes for different values of sensitivity/ specificity claimed by the manufacturer.

Sensitivity/ Specificity	*Sample size for each of the 04 target pathogens[a]: Minimum number of positive samples[¥]*	*Composition of positive samples for each pathogen/type/subtype[#]*	*Total number of positive samples (including all 04 pathogens)*	*Sample size: Minimum number of negative samples[¥]*	*Minimum number of cross reactive* samples among the negative samples*
99%	16 (rounded to 20 for better distribution of samples)	Strong positive = 06 Moderate positive = 07 Weak positive = 07	80	20	5
95%	77 (rounded to 80 for better distribution of samples)	Strong positive = 24 Moderate positive = 28 Weak positive = 28	320	80	20
90%	146 (rounded to 155 for better distribution of samples)	Strong positive = 45 Moderate positive = 55 Weak positive =55	620	150	38
85%	207 (rounded to 215 for better distribution of samples)	Strong positive = 63 Moderate positive = 76 Weak positive = 76	860	210	53
80%	259 (rounded to 260 for better distribution of samples)	Strong positive = 78 Moderate positive = 91 Weak positive = 91	1040	260	65

[a]Influenza A: (H1N1) pdm09, Influenza A/H3N2, Influenza B (Victoria), and SARS CoV-2

[#]Strong positive: (Ct value <25)
Moderate positive: (Ct value between 25-30)
Weak positive: (Ct value >30 and ≤ 34)

[¥] Nasopharyngeal or oropharyngeal swabs in virus transport medium (VTM) to be used

*Samples positive for common respiratory viruses (such as Parainfluenza viruses, Adenoviruses, Rhinoviruses, common human coronaviruses, RSV), other than the ones targeted by the kit under evaluation. Equal distribution of cross-reactive viruses is desirable.

For multiplex assays targeting influenza and SARS-CoV-2, samples positive for these viruses may be suitably interchanged for assessing cross-reactivity.
Influenza B (Yamagata) is not mentioned in the table in view nil/minimal global and national circulation of the strain for the past few years. Samples positive for influenza B (Yamagata) should be used for performance evaluation of kits if the strain starts circulating in the country.

It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

2. Acceptance Criteria for the kit:

Sensitivity for each pathogen/ type/ subtype: ≥95%
Specificity for each pathogen/ type/ subtype: ≥99%
Cross-reactivity: Minimal per pathogen/ type/ subtype
Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥80 positive samples and ≥20 negative samples should be tested for evaluation for each pathogen/ type/ subtype.

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

Protocol E:

Performance Evaluation of Molecular IVD Kit detecting Influenza virus, SARS-CoV-2 and Respiratory Syncytial Virus (RSV) in multiplex format

To assess the performance of multiplex assays, Protocols A, B or C can be used as per kit format to check the performance of each virus for its sensitivity and specificity assessment, including cross reactivity, repeatability, reproducibility and Lot to lot variation.

A comprehensive report can be generated which will include sensitivity and specificity for all targets.

Sample size for multiplex molecular assay (as per the scope outlined in the document) detecting Influenza virus, SARS-CoV-2 and Respiratory Syncytial Virus (RSV) in multiplex format is given below. All other relevant performance criteria (such as repeatability/ reproducibility/ cross-reactivity) outlined in the single plex protocols (Protocol A, B and C) are to be essentially followed and met for each target pathogen/type/subtype.

1. **Sample size and sample panel composition for evaluation of performance characteristics**:

Sample sizes of positive and negative samples against different values of sensitivity and specificity are provided in Table 7. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate ≤5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity. Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the number of positive samples.*
- *n (sp) is the number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*

- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Table 7. Sample sizes for different values of sensitivity/ specificity claimed by the manufacturer.

Sensitivity/ Specificity	*Sample size for each of the 06 target pathogens[a]: Minimum number of positive samples[¥]*	*Composition of positive samples for each pathogen/type/subtype[#]*	*Total number of positive samples (including all 06 pathogens)*	*Sample size: Minimum number of negative samples[¥]*	*Minimum number of cross reactive* samples among the negative samples*
99%	16 (rounded to 20 for better distribution of samples)	Strong positive = 06 Moderate positive = 07 Weak positive = 07	120	20	5
95%	77 (rounded to 80 for better distribution of samples)	Strong positive = 24 Moderate positive = 28 Weak positive = 28	480	80	20
90%	146 (rounded to 155 for better distribution of samples)	Strong positive = 45 Moderate positive = 55 Weak positive =55	930	150	38
85%	207 (rounded to 215 for better distribution of samples)	Strong positive = 63 Moderate positive = 76 Weak positive = 76	1290	210	53
80%	259 (rounded to 260 for better distribution of samples)	Strong positive = 78 Moderate positive = 91 Weak positive = 91	1560	260	65

[a]Influenza A: (H1N1) pdm09, Influenza A/H3N2, Influenza B (Vivtoria), SARS CoV-2, RSV A, and RSV B

[#]Strong positive: (Ct value <25)
Moderate positive: (Ct value between 25-30)
Weak positive: (Ct value >30 and ≤ 34)

[¥] Nasopharyngeal/ oropharyngeal swabs in virus transport medium (VTM) to be used

* Samples positive for common respiratory viruses (such as Parainfluenza viruses, Adenoviruses, Rhinoviruses, common human coronaviruses), other than the ones targeted by the kit under evaluation. Equal distribution of cross-reactive viruses is desirable.

For multiplex assays targeting influenza, SARS-CoV-2, and RSV, samples positive for these viruses may be suitably interchanged for assessing cross-reactivity.

Influenza B (Yamagata) is not mentioned in the table in view nil/minimal global and national circulation of the strain for the past few years. Samples positive for influenza B (Yamagata) should be used for performance evaluation of kits if the strain starts circulating in the country.

It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure

adequate power of the study in case the kit falls short of claimed performance characteristics.

2. Acceptance Criteria for the kit:

Sensitivity for each pathogen/ type/ subtype: ≥95%
Specificity for each pathogen/ type/ subtype: ≥99%
Cross-reactivity: Minimal per pathogen/ type/ subtype
Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥80 positive samples and ≥20 negative samples should be tested for evaluation for each pathogen/ type/ subtype.

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

Reference for sample size calculation:

1. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

Performance Evaluation Report Format

Performance evaluation report for Respiratory Virus in-vitro molecular diagnostic kit

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak/simulated samples)	
	Negative samples (provide details, including cross reactivity panel/simulated samples)	

i. **Analytes/Pathogens targeted by the kit under evaluation:**
……
ii. ……
iii. ……
iv. ……
v. ……
vi. ……
vii. ……

RESULTS INTERPRETATION

SENSITIVITY AND SPECIFICITY FOR INDIVIDUAL VIRUS TARGETS

1. **Sensitivity and specificity for Influenza A (H1N1) pdm09**

Name of the Kit Under Evaluation		Reference assay		
		Positive	**Negative**	**Total**
	Positive			
	Negative			
	Total			

	Estimate (%)	**CI 95%**
Sensitivity		
Specificity		

2. **Sensitivity and specificity for Influenza A (H3N2)**

Name of the Kit Under Evaluation		Reference assay		
		Positive	**Negative**	**Total**
	Positive			
	Negative			
	Total			

	Estimate (%)	**CI 95%**
Sensitivity		
Specificity		

3. **Sensitivity and specificity for Influenza B (Victoria)**

Name of the Kit Under Evaluation		Reference assay		
		Positive	**Negative**	**Total**
	Positive			
	Negative			
	Total			

	Estimate (%)	**CI 95%**

Sensitivity		
Specificity		

4. **Sensitivity and specificity for Influenza B (Yamagata) –to be assessed if the strain circulates nationally and/or globally**

Name of the Kit Under Evaluation		Reference assay		
		Positive	**Negative**	**Total**
	Positive			
	Negative			
	Total			

	Estimate (%)	**CI 95%**
Sensitivity		
Specificity		

5. **Sensitivity and specificity for SARS-CoV-2**

Name of the Kit Under Evaluation		Reference assay		
		Positive	**Negative**	**Total**
	Positive			
	Negative			
	Total			

	Estimate (%)	**CI 95%**
Sensitivity		
Specificity		

6. **Sensitivity and specificity for RSV**

Name of the Kit Under Evaluation		Reference assay		
		Positive	**Negative**	**Total**
	Positive			
	Negative			

	Total			

	Estimate (%)	**CI 95%**
Sensitivity		
Specificity		

7. Sensitivity and specificity for RSV A

Name of the Kit Under Evaluation		Reference assay		
		Positive	**Negative**	**Total**
	Positive			
	Negative			
	Total			

	Estimate (%)	**CI 95%**
Sensitivity		
Specificity		

8. Sensitivity and specificity for RSV B

Name of the Kit Under Evaluation		Reference assay		
		Positive	**Negative**	**Total**
	Positive			
	Negative			
	Total			

	Estimate (%)	**CI 95%**
Sensitivity		
Specificity		

a) **Cross-reactivity Analysis:**
b) **Invalid test rate:**
c) **Repeatability Assessment:**
d) **Precision (Reproducibility):**

Lot to Lot

Details of lots tested (3 lots to be tested):

1. Lot No.: Lot No: Tested By:
2. Lot No.: Lot No: Tested By:
3. Lot No.: Lot No: Tested By:

- **Lot-to-lot variation was <u>observed / not observed.</u>**

a. Internal Control Analysis:

<u>Conclusion:</u> Satisfactory / Not satisfactory

<u>RECOMMENDATIONS:</u>

Suggestions for improvements or modifications (if applicable):

_____

____

- **ICMR-CDSCO guidelines were followed for kit performance evaluation.**

This evaluation report is exclusively for_____ In Vitro Molecular Diagnostic Kit manufactured by _____.

Sensitivity and specificity have been assessed in controlled lab settings using the kits of the Lot number:

i. Lot No._____,
ii. Lot No._____,
iii. Lot No._____,

Provided by the manufacturer, using ……. samples. Results should not be extrapolated to other sample types.

<u>DISCLAIMER:</u>

1. This validation process does not approve/disapprove the Kit design.

2. This validation process does not certify user friendliness of the Kit.

3. Influenza and SARS-CoV-2 are continuously evolving viruses and therefore primer probe sequences of the assay may require periodic updates, which will amount to a changed version of the assay. Re-validation is required for changed version of the assay, and needs to be considered while issuing license

Signature of the Lab Manager **Signature of the Lab Director**

Signature of Head of the Institute

Seal of Head of the Institute

Annexure-1: Information on Operational and Test Performance Characteristics Required from Manufacturers

The manufacturer should provide the following details about the IVD:

1. Instructions for Use
2. Scope of the IVD: to diagnose influenza and/or SARS-CoV-/RSV.
3. Intended Use Statement
4. Principle of the assay
5. Intended testing population (cases of ARI/ILI/SARI)
6. Intended user (laboratory professional and/or health care worker at point-of-care)
7. Lot/batch No.
8. Date of manufacture
9. Date of Expiry
10. Information on operational Characteristics
    - i. Configuration of the kit
    - ii. Requirement of any additional equipment, device
    - iii. Requirement of any additional reagents
    - iv. Operation conditions
    - v. Storage and stability before and after opening
    - vi. Internal control provided or not
    - vii. Quality control and batch testing data
    - viii. Biosafety aspects- waste disposal requirements
11. Information on Test Performance Characteristics
    - i. Type of sample-NP/OP swab, other respiratory specimen
    - ii. Volume of sample
    - iii. Any specific sample NOT to be tested
    - iv. Any additional sample processing required
    - v. Any additional device/consumable like sample transfer device, pipette, tube, etc required
    - vi. Name of analyte to be detected
    - vii. Pathogens targeted by the kit

viii. Time taken for testing

ix. Time for result reading and interpretation

x. Manual or automated(equipment)reading

xi. Limit of detection/Limit of Quantification and range of detection

xii. Diagnostic sensitivity

xiii. Diagnostic specificity

xiv. Stability and reproducibility

xv. Training required for testing

xvi. If yes, duration

xvii. Details of Cut-off and /or Equivocal Zone for interpretation of test

xviii. Interpretation of invalid and indeterminate results to be provided

xix. It is recommended to provide data demonstrating accuracy and precision

xx. Limit of detection

HUMAN METAPNEUMOVIRUS REAL-TIME PCR

List of Contributors:

A. Working Group:

1. Dr. Varsha Potdar, Scientist-E, ICMR-National Institute of Virology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
2. Ms. Krittika Bhattacharyya, Statistical Officer (Planning), Directorate of Economics and Statistics, Government of National Capital Territory of Delhi
3. Dr. Labanya Mukhopadhyay, Scientist-C, Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

B. Review Committee:

1. Dr. Mandeep Chadha, Former Scientist-G, ICMR-National Institute of Virology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
2. Dr. Shobha Broor, Former Professor and Head of Virology Division, All India Institute of Medical Sciences, Delhi, and current Emeritus Professor, SGT University, Gurgaon, Haryana
3. Mr. Pramod Meshram, Deputy Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
4. Dr. Sella Senthil, Assistant Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
5. Dr. Md Omair Anwar, Drugs Inspector (Medical Devices), IVD Division, Central Drugs Standard Control Organization, Government of India
6. Dr. Nivedita Gupta, Scientist-G and Head of the Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

Performance evaluation protocol for Human Metapneumovirus real-time PCR kit

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding in-vitro diagnostic kit (IVD) performance.

This recommendation focuses on the laboratory performance evaluation of Human Metapneumovirus (hMPV) virus real time PCR kit. All clinical samples tested in the study should be evaluated in accordance with the candidate test's instructions for use.

II. Purpose:

To evaluate the performance characteristics of hMPV real-time PCR kits in the diagnosis of hMPV infection/ disease using irreversibly de-identified leftover archived/ spiked clinical samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.

2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should be well-equipped and establish their proficiency through ALL of the following:
   A. Accreditation for at least one Quality management system for at least one respiratory viral pathogen molecular testing (accreditation for Testing Lab /

Calibration Lab as per ISO/IEC 17025, Medical Lab as per ISO 15189, PT provider as per ISO/IEC 17043), or CDSCO approved Reference laboratory.

B. Staff training: All the staff involved in hMPV virus IVD evaluation should undergo hands-on training and competency testing on following

- Preparation & characterization of reference sample panel (at least 2 staff)
- Handling of hMPV RT-PCR kits received for performance evaluation (Verification/Storage/Unpacking etc).
- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of hMPV RNA evaluation panel

A well characterised panel of hMPV positive human samples is a critical requirement for evaluation of these RT-PCR IVD kits. A statistically significant number of clinical samples should be used for the evaluation.

The sample type for hMPV detection is nasopharyngeal/oropharyngeal swab. If a kit claims to detect hMPV across several sample types, attempt should be made to evaluate the assay across all the sample types. In case all the sample types mentioned in the IFU are not available with the lab, the performance evaluation report should clearly mention the sample type against which the kit is evaluated. There should be no ambiguity about the type of sample used for evaluation.

4. RNA extraction

RNA extraction should be performed as per reference test and index test Instructions for Use (IFU).

If the manufacturer of the index test recommends a specific RNA extraction kit/system, the same needs to be provided by the manufacturer if the evaluating lab is unable to procure the same/it is not available within the lab's IVD evaluation scope.

5. Real-Time PCR System

PCR should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Real-time closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.

6. Internal control/Extraction control

The index test must have an internal control (housekeeping gene), with or without an extraction control (RNA added before extraction to a sample).

7. Reference assay:

WHO Pre-Qualified/ US FDA/ ATAGI Austtralia/ PMDA Japan approved real-time PCR assay/ ICMR-NIV Pune in-house Real Time PCR Assay should be used as the Reference Assay.

All positive samples should be confirmed positive by the reference assay.

All negative samples should be confirmed negative by the reference assay.

8. Sample size for performance evaluation: Sample size is calculated assuming 95% sensitivity and specificity of the index test, 95% confidence level, absolute precision of 5% and ≤5% invalid test rate. A minimum of 77 (rounded to 80) positive clinical samples and a minimum of 77 (rounded to 80) negative clinical samples are required for performance evaluation. However, for negative samples, a minimum of 115 specimens are suggested to account for a rigorous cross reactivity panel.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \, x \, S_e \, (1 - S_e)}{d^2 \, x \, (1 - IR)}$$

$$n_{sp} \geq \frac{Z^2 \, x \, S_p \, \left(1 - S_p\right)}{d^2 \, x \, (1 - IR)}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

9. Sample panel composition:

A. Human samples

A.1 Positive samples (Minimum n=80): Clinical samples positive by the reference real-time PCR assay

A.1.1 Strong positive (Ct value <25) = 24 samples

A.1.2. Moderate positive (Ct value between 25-30) = 28 samples

A.1.3 Weak positive (Ct value >30-35) = 28 samples

Note:

If possible, attempt should be made to include all lineages of hMPV in the positive sample panel.

A.2 Negative samples: All negative samples should be negative by reference real-time PCR assay. Distribution of the negative samples should be as follows:

A.2.1 NP/OP swab from individuals with respiratory infection that are negative for hMPV RNA = 30 samples

A.2.2 NP/OP swab from apparently healthy individuals with no respiratory symptoms = 20 samples

A.2.3 Cross reactivity panel (Table 1): Samples negative for hMPV RNA but positive for other common respiratory viruses = 65 samples

Table 1: Cross reactivity panel for performance evaluation of HMPV real time PCR kit

S.N.	Pathogen	Minimum no. of positive samples needed (n=65)	Additional comments
i.	RSV A @	5	In case adequate number of one RSV type is unavailable, supplement with the available RSV type
ii.	RSV B @	5	
iii.	Measles @	5	-
iv.	Mumps @	5	Buccal swab is the preferred sample type for Mumps, and the same (or throat swab) should be used for evaluation
v.	Seasonal Influenza A (H1N1pdm09 **and** H3N2) @	10 (5 of each)	-
vi.	Seasonal Influenza B (Victoria, with/without Yamagata) @	5	-
vii.	SARS-CoV-2	5	-
viii.	Seasonal coronaviruses @	3	OC43 **AND** 229E
ix.	Respiratory Adenovirus @	5	Representation from all respiratory types is desirable

x.	Human Respiroviruses 1 and 3, Human Rubulaviruses 2 and 4 @	5	Representation from all types is desirable
xi.	Rhinovirus @	5	In case samples available with the lab are not typed into Rhinovirus and non-Rhinovirus Enteroviruses, please use 10 such samples to represent these 2 pathogens
xii.	Enterovirus *	5	
xiii.	Cytomegalovirus *	2	Lower respiratory specimen positive for CMV is acceptable

Cross reactivity panel is arranged in descending order of priority.

The pathogens marked @ are essentially to be tested.

*It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*

If available, samples positive for relevant bacterial pathogens and other relevant viruses (with which majority of the population is likely to be infected), should also be included in the cross-reactivity panel.

Commercially available validated standard panels that are accepted by accreditation agencies can also be used if clinical samples with required target for cross-reactivity analysis is not available.

10. Evaluation method:

The index test and the reference assay should be run simultaneously on the sample panel, and results should be recorded.

11. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

12. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.

False negative samples: These are samples positive by reference assay and negative by index test.

13. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

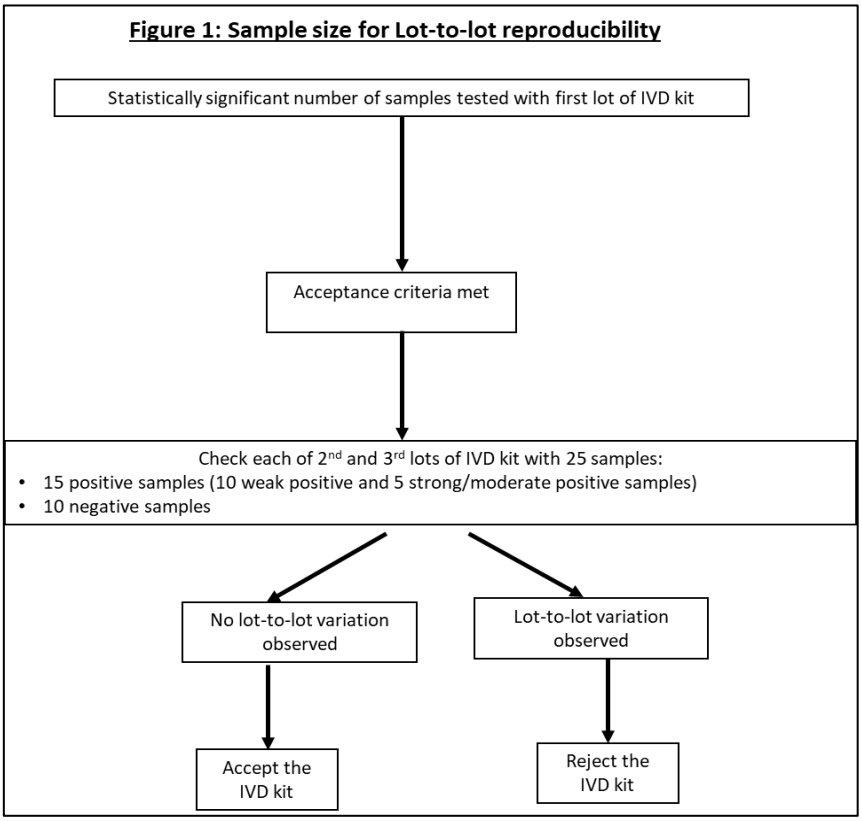
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

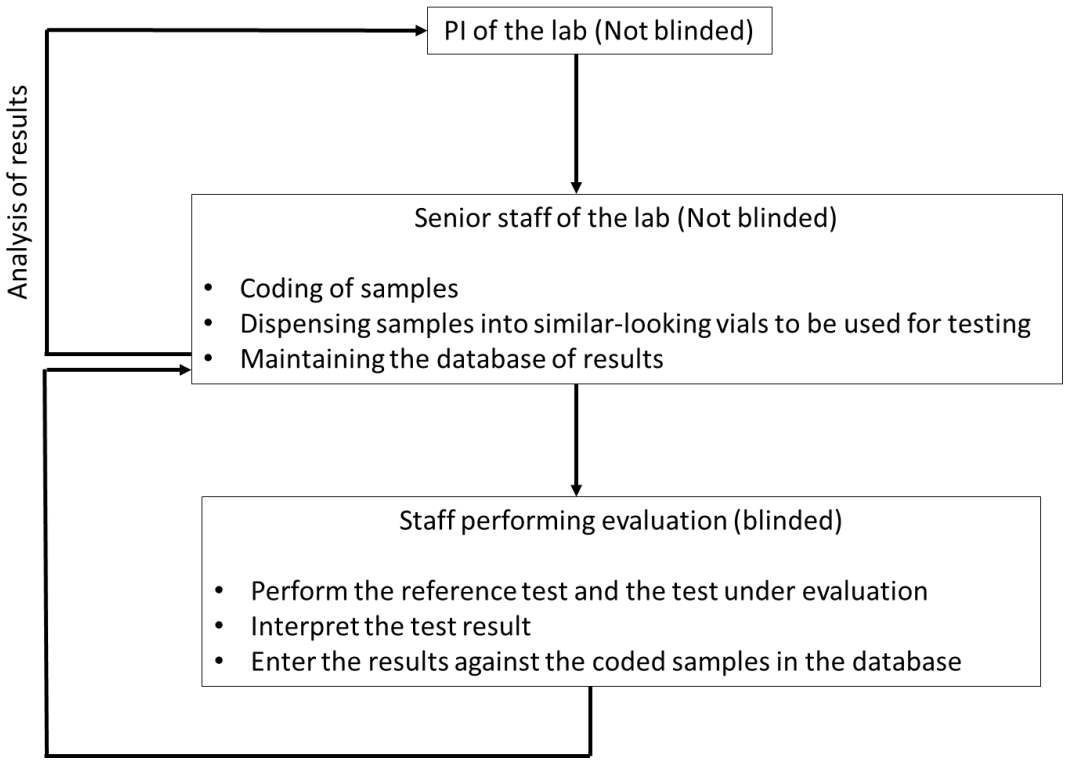
d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report Ct standard deviation across repeats (≤1cycle deviation recommended).

14. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



15. Acceptance Criteria:

Sensitivity: ≥95%

Specificity: ≥98%

Cross reactivity with other viruses as outlined in the negative sample panel: Minimal

Invalid test rate: ≤5%

16. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. U.S. Food and Drug Administration: Testing for Human Metapneumovirus (hMPV) Using Nucleic Acid Assays - Class II Special Controls Guidance for Industry and FDA Staff. 2009. Available at: https://www.fda.gov/medical-devices/guidance-documents-medical-devices-and-radiation-emitting-products/testing-human-metapneumovirus-hmpv-using-nucleic-acid-assays-class-ii-special-controls-guidance#3 [Accessed on January 11, 2025]

2. Amarasinghe, G.K., Ayllón, M.A., Bào, Y. *et al.* Taxonomy of the order *Mononegavirales*: update 2019. *Arch Virol* 164, 1967–1980 (2019). https://doi.org/10.1007/s00705-019-04247-4

VII. Performance evaluation report format

PERFORMANCE EVALUATION REPORT FOR HUMAN METAPNEUMOVIRUS (HMPV) REAL-TIME PCR KITS

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number: Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: clinical/spiked, strong, moderate, weak/simulated samples)	
	Negative samples (provide details (clinical/spiked), including cross reactivity panel/simulated samples)	

Results

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of hMPV virus real-time PCR	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity
- Invalid test rate

- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ….. sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note:

This report is exclusively for Human Metapneumovirus…….. Kit (Lot No……) manufactured by …… (supplied by …….)

The kit has been validated against the pathogen (as a whole) with statistically significant sample size, and NOT against different lineages of the pathogen.

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal …….

Annexure-1: Information on Operational and Test Performance Characteristics Required from Manufacturers

The manufacturer should provide the following details about the IVD:

1. Instructions for Use
2. Scope of the IVD:
3. Intended Use Statement
4. Principle of the assay
5. Intended testing population (cases of ARI/ILI/SARI)
6. Intended user (laboratory professional and/or health care worker at point-of-care)
7. Lot/batch No.
8. Date of manufacture
9. Date of Expiry
10. Information on operational Characteristics
    - i. Configuration of the kit/device
    - ii. Requirement of any additional equipment, device
    - iii. Requirement of any additional reagents
    - iv. Operation conditions
    - v. Storage and stability before and after opening
    - vi. Internal control provided or not
    - vii. Quality control and batch testing data
    - viii. Biosafety aspects- waste disposal requirements
11. Information on Test Performance Characteristics
    - i. Type of sample-NP/OP swab, other respiratory specimen
    - ii. Volume of sample
    - iii. Any specific sample NOT to be tested
    - iv. Any additional sample processing required
    - v. Any additional device/consumable like sample transfer device, pipette, tube, etc required
    - vi. Name of analyte to be detected
    - vii. Pathogens targeted by the kit

viii. Time taken for testing

ix. Time for result reading and interpretation

x. Manual or automated (equipment) reading

xi. Limit of detection/Limit of Quantification and range of detection

xii. Diagnostic sensitivity

xiii. Diagnostic specificity

xiv. Stability and reproducibility

xv. Training required for testing

xvi. If yes, duration

xvii. Details of Cut-off and /or Equivocal Zone for interpretation of test

xviii. Interpretation of invalid and indeterminate results to be provided

xix. It is recommended to provide data demonstrating accuracy and precision

*****Please mention "Not applicable" against sections not pertaining to the kit.

MULTIPLEX RESPIRATORY VIRUS REAL TIME PCR

List of Contributors:

A. Working Group:

1. Dr. Labanya Mukhopadhyay, Scientist-C, Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India
2. Ms. Krittika Bhattacharyya, Statistical Officer (Planning), Directorate of Economics and Statistics, Government of National Capital Territory of Delhi

B. Review Committee:

1. Dr. Mandeep Chadha, Former Scientist-G, ICMR-National Institute of Virology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
2. Dr. Shobha Broor, Former Professor and Head of Virology Division, All India Institute of Medical Sciences, Delhi, and current Emeritus Professor, SGT University, Gurgaon, Haryana
3. Dr. Vasanthapuram Ravi, Former Dean Research and Head of Neurovirology, National Institute of Mental Health and Neuro-Sciences, Bengaluru, Karnataka
4. Mr. Pramod Meshram, Deputy Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
5. Dr. Sella Senthil, Assistant Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
6. Dr. Md Omair Anwar, Drugs Inspector (Medical Devices), IVD Division, Central Drugs Standard Control Organization, Government of India
7. Dr. Nivedita Gupta, Scientist-G and Head of the Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

Performance evaluation protocol for multiplex respiratory virus real-time PCR kit

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured diagnostic kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

This recommendation focuses on the laboratory performance evaluation of multiplex respiratory virus real time PCR kit. All clinical samples tested in the study should be evaluated in accordance with the candidate test's instructions for use.

II. Purpose:

To evaluate the performance characteristics of multiplex respiratory virus real-time PCR kits using irreversibly de-identified leftover archived clinical/spiked samples.

III. Scope of the document:

This document outlines performance evaluation protocol for multiplex real time PCR assays detecting the following respiratory viruses of utmost importance in human clinical specimens (Table 1), as determined by ICMR appointed working group and expert group of physicians and clinical microbiologists following extensive literature review and real-life experience. This pathogen list has been developed as part of the National One Health Mission.

Table 1: List of respiratory viruses within the scope of this performance evaluation protocol

1. Influenza virus A
2. Influenza virus B
3. SARS Coronavirus-2
4. Respiratory syncytial virus
5. Adenovirus
6. Human Respiroviruses 1 and 3 and Human Rubulaviruses 2 and 4 (erstwhile Human Parainfluenzaviruses 1-4)
7. Human metapneumovirus
8. Measles virus
9. Rhinovirus
10. Human Bocavirus
11. Enterovirus
12. Cytomegalovirus

IV. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.

2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

V. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

VI. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified archived/ spiked clinical samples

2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should be well-equipped and establish their proficiency through ALL of the following:

   A. Accreditation at least one of the Quality management systems for at least one respiratory viral pathogen molecular testing (accreditation for Testing Lab / Calibration Lab as per ISO/IEC 17025, Medical Lab as per ISO 15189, PT provider as per ISO/IEC 17043), or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD evaluation should undergo hands-on training and competency testing on the following:

      - Preparation & characterization of reference sample panel
      - Handling of multiplex respiratory virus RT-PCR kits received for performance evaluation (Verification/Storage/Unpacking etc).
      - Testing
      - Data handling, data safety & confidentiality

3. Preparation of multiplex respiratory virus evaluation panel:

A well characterised panel of positive and negative clinical samples is a critical requirement for evaluation of these RT-PCR IVD kits. Also, a statistically significant number of clinical samples should be used for the evaluation.

The sample type for respiratory virus detection is usually nasopharyngeal/oropharyngeal swab. If a kit claims to detect these viruses across several sample types, attempt should be made to evaluate the assay across all the sample types. In case all the sample types mentioned in the IFU are not available with the lab, the performance evaluation report should clearly

mention the sample type against which the kit is evaluated. There should be no ambiguity about the type of sample used for evaluation.

4. Nucleic acid extraction:

Nucleic acid extraction should be performed as per reference test and index test Instructions for Use (IFU).

If the manufacturer of the index test recommends a specific RNA extraction kit/system, the same needs to be provided by the manufacturer if the evaluating lab is unable to procure the same/it is not available within the lab's IVD evaluation scope.

**Caution is advised in the selection of a nucleic acid extraction kit since the target pathogens comprise both RNA and DNA viruses.*

5. Real-Time PCR System:

PCR should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Real-time closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.

6. Internal control/Extraction control:

The test under evaluation (index test) must have an internal control (housekeeping gene), with or without an extraction control (nucleic acid added before extraction to a sample).

7. Reference assay:

The following points are to be noted:

i. A WHO Pre-Qualified/ US FDA/ ATAGI Australia/ PMDA Japan approved single plex (for a particular target pathogen) or multiplex real-time PCR assay/ ICMR-NIV Pune in-house single plex (for a particular target pathogen) or multiplex Real Time PCR Assay should be used as the reference assay for each target pathogen/type/subtype.
ii. Since the list of target pathogens is extensive, a combination of single plex and/or multiplex assays may be used as the reference assay(s), as long as these reference assays satisfy the criteria outlined in point 7(i).

All samples positive for a particular pathogen should be confirmed positive by the reference assay.

All samples negative for a particular pathogen should be confirmed negative by the reference assay.

8. Sample size for performance evaluation: The 2009 FDA guidance document "*Respiratory Viral Panel Multiplex Nucleic Acid Assay - Class II Special Controls Guidance for Industry and FDA Staff*", recommends including a sufficient number of prospectively collected samples for each specimen type to generate a result with at least 90% sensitivity with a lower bound of the two-sided 95% confidence interval (CI) greater than 80, and demonstrate specificity with a lower bound of the two-sided 95% CI greater

than 90%. In accordance with these guidelines and for feasibility of evaluation of these extensive multiplex panels, sample size for each pathogen is calculated assuming ≥90% sensitivity and specificity of the index test, 95% confidence level, absolute precision of 7.5%, and ≤5% invalid test rate. A minimum of 65 positive clinical samples (rounded to 70) and a minimum of 65 negative clinical samples for each target pathogen are required for performance evaluation of the assay. However, 120 negative samples are recommended per pathogen to account for an extensive cross reactivity panel. Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

The details of sample requirement are outlined in Table 2.

Table 2: No. of samples required for performance evaluation:

Pathogen	Minimum no. of positive samples needed per pathogen	Minimum no. of negative samples recommended per pathogen
1. Influenza virus A*	70	120
2. Influenza virus B*	70	120
3. SARS Coronavirus-2	70	120
4. Respiratory syncytial virus*	70	120
5. Adenovirus*	70	120
6. Human Respirovirus 1 and Human Respirovirus 3 and Human Rubulavirus 2 and Human Rubulavirus 4*	70	120
7. Human metapneumovirus *	70	120
8. Measles virus	70	120
9. Rhinovirus**	70	120

10. Human Bocavirus	70	120
11. Enterovirus**	70	120
12. Cytomegalovirus	70	120

***If a kit claims to differentiate between virus types/subtypes, please use minimum 70 positive samples and minimum 120 negative samples for each virus type/subtype. If such type/subtype specific samples are not available (only for predicate device) or if the kit does not claim to differentiate between pathogen types/subtypes, and the kit is evaluated against the pathogen as a whole, the reports should be issued with a disclaimer that performance characteristics against pathogen types/subtypes have not been evaluated separately. However, in such a scenario, the evaluating centre should try to include all types/subtypes of the pathogen in the evaluation panel (even if the numbers are not statistically significant for each pathogen type).**

****If clinical samples positive separately for Rhinovirus/Enterovirus are not available (only for predicate device), or if the kit does not differentiate between Enteroviruses and Rhinoviruses, please use minimum 70 samples positive for Rhinovirus/Enterovirus in the positive sample panel and issue the reports with a disclaimer that performance characteristics against Rhinovirus/Enterovirus have not been evaluated separately.**

Influenza virus, SARS Coronavirus 2, Respiratory Syncytial Virus and Human Metapneumovirus positive samples used for evaluation should have been collected within the past 1 year.

Notes for Table 2:

1. Samples positive for currently circulating virus strains should be used in the positive sample panel, with representation from all virus types/subtypes.
2. Sample positive for a particular virus type and negative for the target pathogen being considered may be used in the negative sample panel for the target pathogen, e.g.: a sample positive for SARS-CoV-2 may be used as a negative sample for RSV.

9. Sample panel composition:

A. Human samples

A.1 Positive samples for each pathogen/ type or subtype of pathogen (Minimum n=70): Clinical samples positive by the reference real-time PCR assay should be included, as per the following criteria

A.1.1 Strong positive (Ct value <25) = 20 samples

A.1.2. Moderate positive (Ct value between 25-30) = 25 samples

A.1.3 Weak positive (Ct value >30-36) = 25 samples

A.2 Negative samples for each pathogen/ type or subtype of pathogen (Minimum n=120): All negative samples should be negative for the target pathogen/ its type or subtype by the reference real-time PCR assay. Distribution of the negative samples should be as follows:

A.2.1 NP/OP swab from individuals with respiratory infection that are negative for the target pathogen/its type or subtype = 35 samples **

A.2.2 NP/OP swab from apparently healthy individuals with no respiratory symptoms = 23 samples **

A.2.3 Cross reactivity panel (Table 3): Samples negative for the target pathogen but positive for other common respiratory viruses = 62 samples ***

Archived frozen sample aliquots if used for the evaluation, should not be thawed more than once.

** *If samples are available with the evaluating lab that satisfy these criteria and are negative for all the pathogens targeted by the kit, the same samples may be included in the negative sample panel for all target pathogens to prevent wastage of resources.*

*** *Same positive samples may be included in the cross-reactivity panel of several target pathogens to prevent wastage of resources e.g.: the same Influenza A virus positive sample may be included in the cross-reactivity panel for RSV, Human Metapneumovirus, SARS-CoV-2 etc.*

Table 3: Cross reactivity panel for performance evaluation of multiplex respiratory virus real time PCR kit

Target Pathogen	**Virus-wise no. of samples needed for cross reactivity analysis**														Total no. of cross reactive samples per pathogen
	Influenza virus A *****	Influenza virus B*****	SARS Coronavirus-2 *****	Respiratory syncytial virus *****	Adenovirus **@**	Human Respiroviruses 1 and 3, Human Rubulaviruses 2 and 4 **#**	Human metapneumovirus**@**	Measles virus *****	Rhinovirus **@ $**	Human Bocavirus	Enterovirus **$**	Cytomegalovirus**.**	Seasonal coronaviruses *****	Rubella	
1. Influenza virus A	0	5	5	5	5	5	5	5	5	5	5	5	5	2	62
2. Influenza virus B	5	0	5	5	5	5	5	5	5	5	5	5	5	2	62
3. SARS Coronavirus-2	5	5	0	5	5	5	5	5	5	5	5	5	5	2	62
4. Respiratory syncytial virus	5	5	5	0	5	5	5	5	5	5	5	5	5	2	62
5. Adeno virus	5	5	5	5	0	5	5	5	5	5	5	5	5	2	62
6. Human Respiroviruses 1 and 3, Human Rubulaviruses 2 and 4	5	5	5	5	5	0	5	5	5	5	5	5	5	2	62
7. Human metapneumovirus	5	5	5	5	5	5	0	5	5	5	5	5	5	2	62
8. Measles virus	5	5	5	5	5	5	5	0	5	5	5	5	5	2	62
9. Rhinovirus	5	5	5	5	5	5	5	5	0	5	5	5	5	2	62
10. Human Bocavirus	5	5	5	5	5	5	5	5	5	0	5	5	5	2	62
11. Enterovirus	5	5	5	5	5	5	5	5	5	5	0	5	5	2	62
12. Cytomegalovirus	5	5	5	5	5	5	5	5	5	5	5	0	5	2	62

*****Include all currently circulating strains/types/subtypes

@It is desirable to have representation from all types of the pathogen, since even approved assays may not always differentiate between pathogen types.

Include at least 1 of each

$ If clinical samples positive separately for Rhinovirus/Enterovirus are not available, please use total 10 samples positive for Rhinovirus/Enterovirus in the cross-reactivity panel for remaining pathogens.

. Can use lower respiratory tract specimen

If a kit claims to differentiate between virus types/subtypes, please use 5 positive samples for each virus type in the cross reactivity panel for other target pathogens. If such type specific samples are not available and the kit is evaluated against the pathogen as a whole, it should be clearly mentioned in the report.

If available, samples positive for relevant bacterial pathogens and other relevant viruses (with which majority of the population is likely to be infected), should also be included in the cross-reactivity panel.

Influenza virus, SARS Coronavirus 2, Respiratory Syncytial Virus and Human Metapneumovirus positive samples used for evaluation should have been collected within the past 1 year.

B. Contrived samples:

Contrived positive and negative samples may be used for evaluation in case of paucity/unavailability of human clinical samples. Positive contrived samples should be positive and negative contrived samples should be negative for the target pathogen/type/subtype using the reference assay. The number and distribution of positive and negative samples, including the cross reactivity panel, should remain the same.

Contrived positive samples (as part of positive sample panel/ cross-reactivity panel) should be prepared by spiking a sample matrix negative for the pathogen with a pathogen-infected cell line, genomic DNA plasmids or RNA transcripts.

It is recommended to demonstrate equivalence between contrived and clinical specimens. Serial dilutions of clinical sample and serial dilutions of contrived sample with targeted levels of analyte should be compared for demonstrating equivalence.

10. Evaluation method:

The index test and the reference assay should be run simultaneously on the sample panel, and results should be recorded.

11. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

12. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

13. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

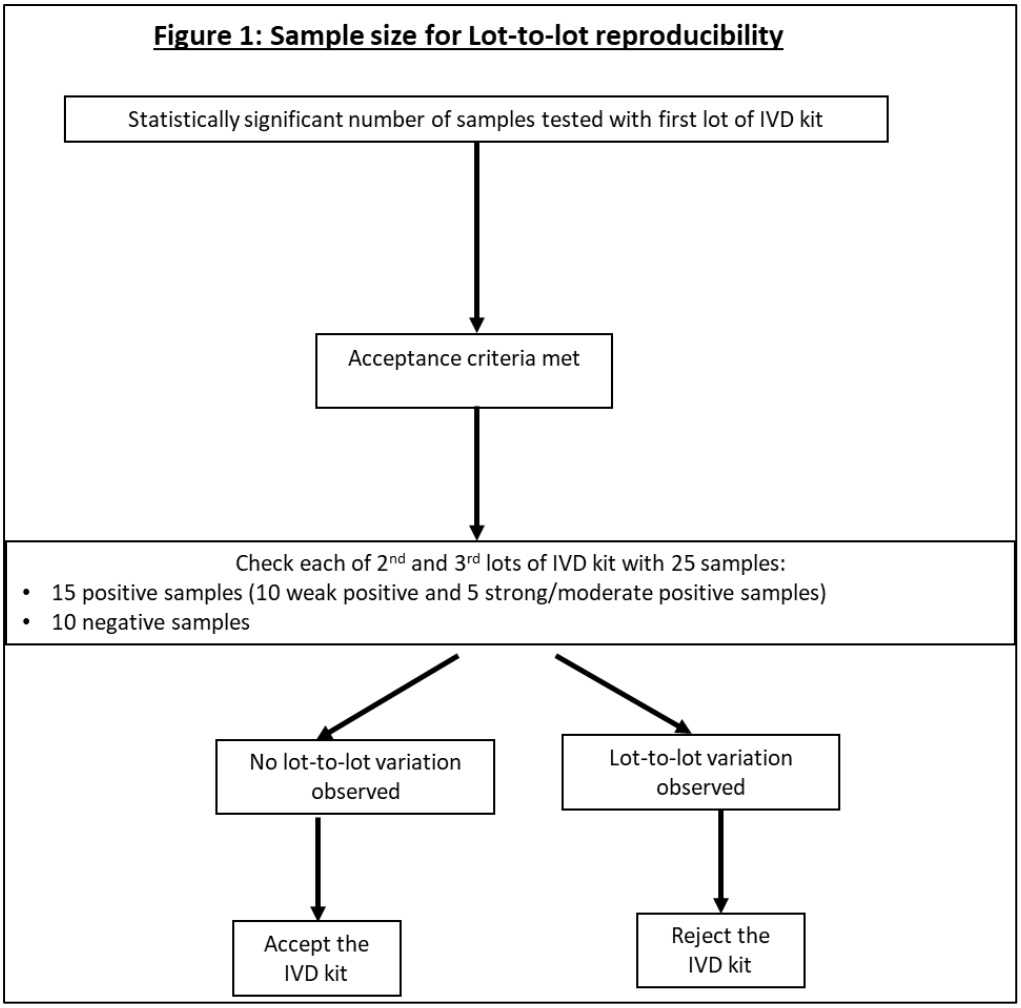
B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility per target should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples per target should be tested 5 times in independent runs.

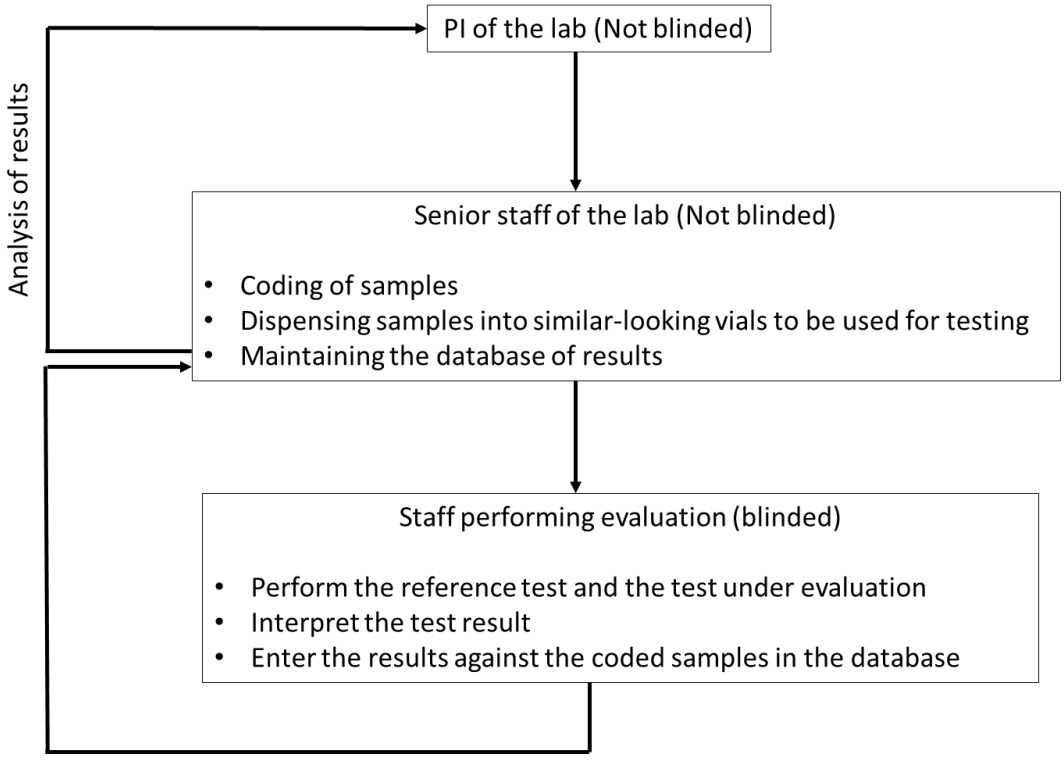
Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report Ct standard deviation across repeats (≤1cycle deviation recommended).

14. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



15. Acceptance Criteria

Sensitivity for each pathogen/type/subtype: ≥90%

Specificity for each pathogen/type/subtype: ≥95%

Cross reactivity with other viruses as outlined in the negative sample panel: Minimal

Invalid test rate: ≤5%

16. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VII. References:

1. Food and Drug Administration. Respiratory Viral Panel Multiplex Nucleic Acid Assay - Class II Special Controls Guidance for Industry and FDA Staff. Available at: https://www.fda.gov/medical-devices/guidance-documents-medical-devices-and-radiation-emitting-products/respiratory-viral-panel-multiplex-nucleic-acid-assay-class-ii-special-controls-guidance-industry-and [Accessed on 22nd January, 2025].
2. Food and Drug Administration. 510(k) Substantial Equivalence Determination Decision Summary, Biofire Diagnostics LLC, FilmArray Pneumonia Panel. Available at: https://www.accessdata.fda.gov/cdrh_docs/reviews/K180966.pdf [Accessed on 19th January 2025]
3. Food and Drug Administration: Testing for Human Metapneumovirus (hMPV) Using Nucleic Acid Assays - Class II Special Controls Guidance for Industry and FDA Staff. 2009. Available at: https://www.fda.gov/medical-devices/guidance-documents-medical-devices-and-radiation-emitting-products/testing-human-metapneumovirus-hmpv-using-nucleic-acid-assays-class-ii-special-controls-guidance#3 [Accessed on January 11, 2025]

VIII. Performance evaluation report format

Performance evaluation report for multiplex respiratory virus real-time PCR kits

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Pathogens detected by the assay		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: clinical/spiked, strong, moderate, weak/simulated samples)	
	Negative samples (provide details: clinical/spiked, including cross reactivity panel/simulated samples)	

Results: Tables 1 and 2 should be made for each pathogen/type of pathogen targeted by the kit under evaluation

Table 1: 2x2 table for sensitivity and specificity calculation (prepare 1 table for each target pathogen /type/ subtype)

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of kit and name of target pathogen	Positive			
	Negative			
	Total			

Table 2: Sensitivity and specificity (calculate for each target pathogen /type/ subtype)

	Estimate (%)	95% CI

Sensitivity		
Specificity		

- Cross reactivity
- Invalid test rate
- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ..... sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note:

This report is exclusively for Human Metapneumovirus…….. Kit (Lot No……) manufactured by …… (supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal …….

Annexure-1: Information on Operational and Test Performance Characteristics Required from Manufacturers

1. The manufacturer should provide the following details about the IVD:
2. Instructions for Use
3. Scope of the IVD:
4. Pathogens/type/subtype of pathogens targeted by the kit
5. Intended Use Statement
6. Principle of the assay
7. Intended testing population (cases of ARI/ILI/SARI)
8. Intended user (laboratory professional and/or health care worker at point-of-care)
9. Lot/batch No.
10. Date of manufacture
11. Date of Expiry
12. Information on operational Characteristics

    i. Configuration of the kit/device

    ii. Requirement of any additional equipment, device

    iii. Requirement of any additional reagents

    iv. Operation conditions

    v. Storage and stability before and after opening

    vi. Internal control provided or not

    vii. Quality control and batch testing data

    viii. Biosafety aspects- waste disposal requirements

13. Information on Test Performance Characteristics

    i. Type of sample-NP/OP swab, other respiratory specimen

    ii. Volume of sample

    iii. Any specific sample NOT to be tested

    iv. Any additional sample processing required

    v. Any additional device/consumable like sample transfer device, pipette, tube, etc required

    vi. Name of analyte to be detected

    vii. Pathogens targeted by the kit

viii. Time taken for testing

ix. Time for result reading and interpretation

x. Manual or automated (equipment) reading

xi. Limit of detection/Limit of Quantification and range of detection

xii. Diagnostic sensitivity

xiii. Diagnostic specificity

xiv. Stability and reproducibility

xv. Training required for testing

xvi. If yes, duration

xvii. Details of Cut-off and /or Equivocal Zone for interpretation of test

xviii. Interpretation of invalid and indeterminate results to be provided

xix. It is recommended to provide data demonstrating accuracy and precision

xx. Limit of detection

*****Please mention "Not applicable" against sections not pertaining to the kit.

MALARIA IN-VITRO DIAGNOSTICS

List of Contributors:

A. Working Group:

1. Dr. Praveen K. Bharti, Scientist-F, ICMR-National Institute of Malaria Research, Department of Health Research, Ministry of Health and Family Welfare, Government of India
2. Dr. Gauri Misra, Scientist-I, Head Molecular Diagnostics and COVID-19 Kit Testing Laboratories, National Institute of Biologicals, Noida
3. Dr Shafeeq K Shahul Hameed, Former Scientific Officer, Center for Brain Research, Indian Institute of Science, Bengaluru, and currently Senior Epidemiologist, Health Emergency Department, Ministry of Public Health, Qatar
4. Dr. Vishal Deo, Scientist-C, ICMR- National Institute for Research in Digital Health and Data Science, Department of Health Research, Ministry of Health and Family Welfare, Government of India
5. Dr. Labanya Mukhopadhyay, Scientist-C, Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

B. Review Committee:

1. Dr. Ashis Das, Senior Professor, Department of Biological Sciences, Birla Institute of Technology and Science, Pilani
2. Dr. Pawan Malhotra, Department of Malaria Biology, International Center for Genetic Engineering & Biotechnology, New Delhi
3. Dr. Lokesh Kori, Scientist-D, Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India
4. Mr. Pramod Meshram, Deputy Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
5. Dr. Sella Senthil, Assistant Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
6. Dr. Md Omair Anwar, Drugs Inspector (Medical Devices), IVD Division, Central Drugs Standard Control Organization, Government of India
7. Dr. Nivedita Gupta, Scientist-G and Head of the Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

Performance evaluation protocol for Malaria Rapid diagnostic test (RDT) kits

I. Background:

CDSCO/ICMR, New Delhi have aimed to facilitate the evaluation and supply of Quality-Assured in vitro Diagnostics (IVD) kits suitable for use in India. Hence, the following guidelines shall establish the uniformity during the performance evaluation of IVD kits The objective of performance evaluation is to independently validate the manufacturer's claim regarding in-vitro diagnostic kit (IVD) performance.

II. Purpose:

To evaluate the performance characteristics of rapid diagnostic test kit for the diagnosis of malaria parasite using irreversibly de-identified leftover archived clinical samples.

III. Requirements:

a) Instructions for use (IFU)

b) Supply of RDT kits under evaluation (with batch no.; lot no.; manufacturing and expiry date and other required details).

c) Evaluation sites/laboratories (With required equipment)

d) Reference test kits

e) Characterised Evaluation panel

f) Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/leftover samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through:

   a) Laboratory accreditation: Accreditation for at least one of the Quality management systems (accreditation for Testing Lab / Calibration Lab (ISO/IEC: 17025),

Medical Lab (ISO: 15189), PT provider (ISO/IEC: 17043) or CDSCO approved Reference laboratory.

b) It is recommended that malaria Medical Device Testing Labs (MDTLs) participate in Quality Control exercises such as EQAP (External Quality Assurance Programme).

c) **Staff training:** All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on the following at suitable malaria labs before initiation of MDTL activity:

- Preparation and characterization of evaluation panel for the respective IVD kit.
- Management of RDT kits (specific for *Plasmodium falciparum / Plasmodium vivax*) received for performance evaluation (Verification/Storage/Unpacking etc.).
- Perform tests interpretation and documentation of results, and reporting.
- Data management and safety and confidentiality.

3. Preparation of sample panel for Malaria RDT kit evaluation

To evaluate the performance of IVD kit, a well characterized species-specific malaria antigen sample panel is required. Statistically significant number of blood samples as defined in this protocol should be used (as mentioned in Table 1). The panel should comprise positive and negative samples as described in section 5.

The reference sample panel should be stored in appropriate storage conditions (depending on the sample type and planned storage duration), and the quality of the panel should be checked periodically (at least once a year) with appropriate tests (e.g.: reference test/parasite culture/enzymatic activity/other relevant test).

4. Reference assay:

WHO Pre-Qualified/ US FDA/ ATAGI Australia/ PMDA Japan approved RDT should be used as reference standard. Parasitemia level should be determined by microscopy and/or other relevant test results.

All positive samples should be confirmed positive by the reference assay.

All negative samples should be confirmed negative by the reference assay.

5. Sample size and sample panel composition for performance evaluation:

Sample sizes of positive and negative samples of each species targeted by the kit against different values of sensitivity and specificity are provided in Tables 1 and 2, with recommended composition. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate of 5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in

the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table).

For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity.

Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Table 1. Positive sample sizes (per species) and composition for different values of sensitivity claimed by the manufacturer for evaluation of Pf (single/combo RDT) or Pv (single/combo RDT)

Sensitivity	*Sample size: Minimum number of positive samples #*	*Composition of positive sample panel*
99%	16 (rounded to 20)	Equal number of high and low parasitemic samples
95%	77 (rounded to 80)	Equal number of high and low parasitemic samples
90%	146 (rounded to 150)	Equal number of high and low parasitemic samples
85%	207 (rounded to 210)	Equal number of high and low parasitemic samples
80%	259 (rounded to 260)	Equal number of high and low parasitemic samples
75%	305 (rounded to 310)	Equal number of high and low parasitemic samples

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Negative sample sizes and composition for different values of specificity claimed by the manufacturer for evaluation of Pf (single/combo RDT) or Pv (single/combo RDT)

Specificity	*Sample size: Minimum number of negative samples #*	*Composition of negative samples[#]*
99%	16 (rounded to 20)	Dengue NS1/IgM positive samples: 03 Chikungunya IgM positive samples:03 Serum reactive for RA factor – low positive and high positive:02 Serum reactive for TPHA/other specific test for syphilis:02 Healthy controls from endemic regions: 10
95%	77 (rounded to 80)	Dengue NS1/IgM positive samples: 10 Chikungunya IgM positive samples:10 Serum reactive for RA factor – low positive and high positive:10 Serum reactive for TPHA/other specific test for syphilis:10 Healthy controls from endemic regions: 40
90%	146 (rounded to 150)	Dengue NS1/IgM positive samples: 18 Chikungunya IgM positive samples:18 Serum reactive for RA factor – low positive and high positive:18 Serum reactive for TPHA/other specific test for syphilis:18 Healthy controls from endemic regions: 78
85%	207 (rounded to 210)	Dengue NS1/IgM positive samples: 26 Chikungunya IgM positive samples:26 Serum reactive for RA factor – low positive and high positive:26 Serum reactive for TPHA/other specific test for syphilis:26 Healthy controls from endemic regions: 106
80%	259 (rounded to 260)	Dengue NS1/IgM positive samples: 35 Chikungunya IgM positive samples:35 Serum reactive for RA factor – low positive and high positive:30 Serum reactive for TPHA/other specific test for syphilis:30 Healthy controls from endemic regions: 130

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Sample panel composition:

Positive samples: Malaria RDT/microscopy positive samples should be obtained from health facilities (*tertiary care centers and their linked hospitals, private clinics, field practice areas etc.*) and confirmed using PCR (Standardized Snounou protocol/US FDA/ATAGI Australia/PMDA Japan approved /WHO Pre-Qualified assay).

- RDT/Microscopy **AND** PCR confirmed positive malaria samples should be characterized further for parasite load on in-house calibrated equipment using blood smear microscopy (and/or other relevant test results).
- Range of Parasitemia: Panel members should have low (≤200 parasites per microliter) to high (≥2000 parasites per microliter) range of Plasmodium falciparum, P. vivax and/or other Plasmodium species, as obtained from microscopy and/or other relevant test results results. Characterized panels **must** contain equal number of samples of both low and high parasitemia.

Note for additional characterization (not mandatory):

*If ELISA is used for characterization of samples in addition to the above-mentioned mandatory tests (RDT/microscopy **AND** PCR), consistent ELISA quantification results should be obtained in ≥3 runs of ELISA experiments performed for each of the three common antigens (PfHRP2, LDH and aldolase), with the results obtained at the 200 p/µL and the 2,000 p/µL being consistent with each other as well (factor of roughly 10 between results). The limit of detection of PfHRP2 is 0.6-74 ng/ mL, PvLDH is 1.6-47.9 ng/ mL, PfLDH is 0.2-53.5 ng/mL, and Pf aldolase is 0-9.9 ng/mL.*

******* If the pool of samples available for testing is sufficiently large in numbers, then the antigen concentration range at the 200 p/µL dilution should be restricted to 5-9.5 ng/mL for PfHRP2, 15-47.9 ng/mL for PvLDH, 10.8-53.5 ng/mL for PfLDH, and 1.7-15ng/mL for Pf aldolase.*

If ELISA for other antigen/antibody is carried out for further characterization (for which limits of detection have not been established), testing should be carried out on calibrated equipment in ≥3 runs of ELISA, leading to their classification as low and high parasitemic samples or strong/moderate/weakly reactive samples.

Negative samples: These are samples negative by RDT/microscopy **AND** PCR(Standardized Snounou protocol/US FDA/ATAGI Australia/PMDA Japan approved /WHO Pre-Qualified assay).

6. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte(s) using the kit under evaluation.

3 positive samples (comprising low parasitemic AND high parasitemic samples) and 3 negative samples per species should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

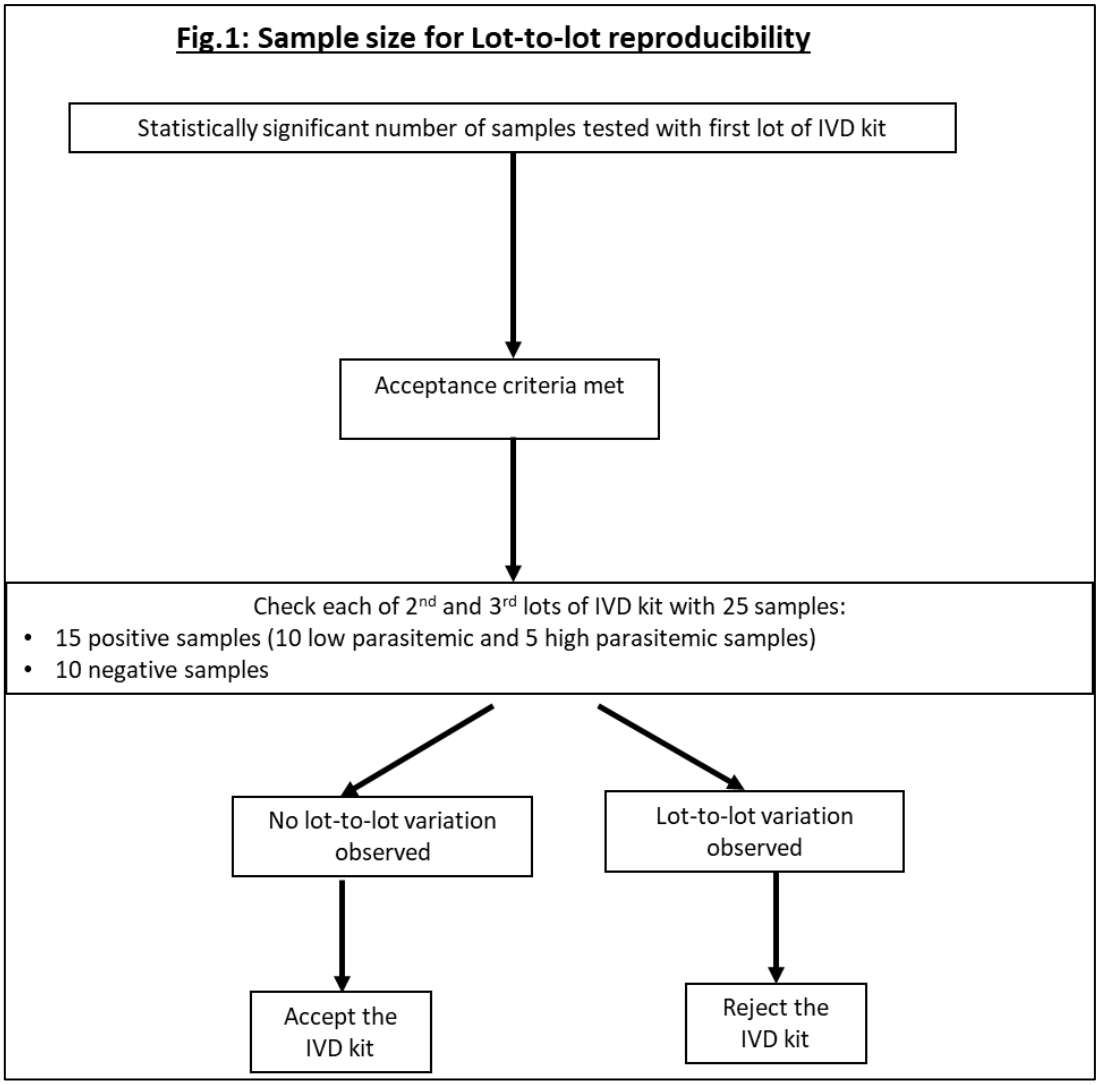
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples per species (15 positive samples comprising 10 low parasitemic AND 5 high parasitemic samples per species, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples per species (15 positive samples comprising 10 low parasitemic AND 5 high parasitemic samples per species, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Fig.1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (comprising low

parasitemic AND high parasitemic samples) and 3 negative samples per species should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (comprising low parasitemic AND high parasitemic samples) and 3 negative samples per species should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (comprising low parasitemic AND high parasitemic samples) and 3 negative samples per species should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

Note: Testing Methodology

Read the instructions for use (IFU) thoroughly. Take out the required number of RDTs kits from the manufacturer-recommended storage conditions. Bring RDTs to room temperature (20°C - 30°C) and thaw the required number of sample aliquots for a minimum of 20 minutes to maximum 60 minutes before performing the test. Note that more than one aliquot may be needed for the testing of each sample. Record the results of the performance evaluation on the recommended report format.

7. Evaluation method:

The reference assay and the index test should be run on the sample panel in parallel.

8. Interpretation of results:

Results should be interpreted as per the IFU of the reference assay and the index test.

9. Resolution of discrepant results:

True positive samples: These are samples positive by both reference assay and index test.
True negative samples: These are samples negative by both reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

10. Acceptance criteria[1]:

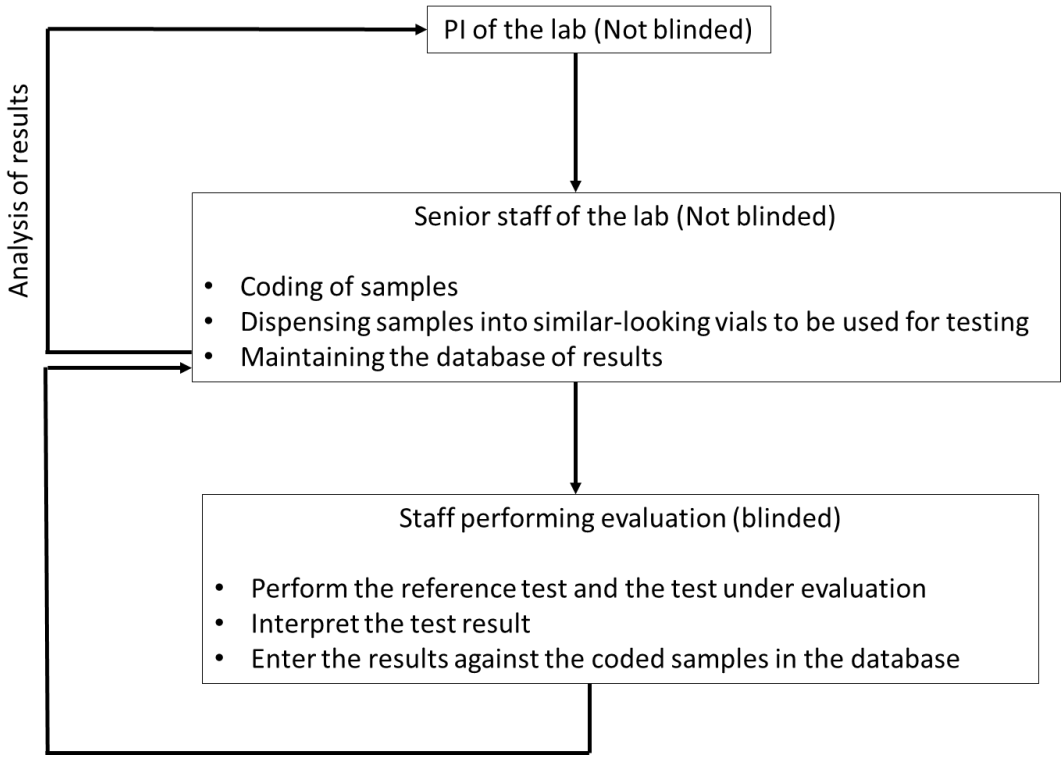
Sensitivity: ≥75% for P. vivax and ≥95% for P. falciparum
Specificity: ≥90% for P. vivax and ≥95% for P. falciparum
Cross-reactivity: Minimal
Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria, ≥310 positive samples and ≥150 negative samples should be tested for *P vivax*, and ≥80 positive samples and ≥80 negative samples should be tested for *P falciparum*.

11. Blinding of laboratory staff

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



12. Publication Rights

The PI(s) of the evaluating labs shall retain publication rights to the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable. Any request of re-validation from the same manufacturer for the same test type will only be entertained if valid proof of change in the kit composition is submitted.

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Ministry of Health and Family Welfare. Guidelines for Bivalent RDT. Available at: guidelines-for-bivalent-rdt.pdf (mohfw.gov.in)
2. World Health Organization. Malaria Rapid Diagnostic Test Performance - Results of WHO product testing of malaria RDTs: round 8 (2016–2018): Available at: https://iris.who.int/bitstream/handle/10665/276190/9789241514965-eng.pdf?sequence=1
3. Snounou G, Viriyakosol S, Zhu XP, Jarra W, Pinheiro L, Do Rosario VE, et al. High sensitivity of detection of human malaria parasites by the use of nested polymerase chain reaction. Molecular and Biochemical Parasitology. 1993;61:315–20.
4. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w
5. World Health Organization. Methods manual for laboratory quality control testing of malaria rapid diagnostic test. Version Ten; 2023

VII. Performance evaluation report format

REPORT FORMAT

Name of the Laboratory

Name of the Institute, (with station)

Certificate of Analysis

File No.: _____

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number: Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: low/high parasitemia/simulated samples)	

	Negative samples (provide details, including cross reactivity panel/simulated samples)	

Results:

		Reference assay …….…… (name)		
		Positive	Negative	Total
Name of index malaria RDT	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity
- Invalid test rate
- Performance: **Satisfactory / Not Satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting on …. samples only, using kits provided by the manufacturer from the batch mentioned above. Results should not be extrapolated for any other sample type.)

Disclaimers

1. This validation process does not approve / disapprove the kit design

2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for …….. Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……

Performance evaluation protocol for Malaria ELISA kits

I. Background:

CDSCO/ICMR, New Delhi, have aimed to facilitate the evaluation and supply of Quality-Assured In Vitro Diagnostics kits suitable for use in India. Hence, the following guidelines shall establish the uniformity during the performance evaluation of IVD kits . The objective of performance evaluation is to independently validate the manufacturer's claim regarding in-vitro diagnostic kit (IVD) performance.

II. Purpose:

To evaluate the performance characteristics of malaria ELISA kits for the diagnosis of malaria parasite infection using irreversibly de-identified leftover archived clinical samples. The malaria ELISA kits are designed to detect antigens (hrp2, LDH, aldolases) occurring in subjects infected with species specific (*P. falciparum, P. vivax*) and stage specific antibodies (MSP1, MSP3, CSP, EBA175 etc.- parasite markers for the purpose of sero-survey).

III. Requirements:

1. Instructions for use (IFU)
2. Supply of ELISA kits under evaluation (with batch no./lot no. expiry date & required details). In case the kit to be evaluated is designed to work in a closed system format, the manufacturer needs to supply the required equipment.
3. Evaluation sites/laboratories (With required equipment)
4. Reference test kits
5. Characterised Evaluation panel
6. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified ELISA kit evaluation laboratories should establish their proficiency through

   a) Laboratory accreditation: Accreditation for at least one of the Quality management systems (accreditation for Testing Lab / Calibration Lab (ISO/IEC: 17025),

Medical Lab (ISO:15189), PT provider ISO/IEC: 17043 or CDSCO approved Reference laboratory.

b) It is recommended that malaria Medical Device Testing Labs (MDTLs) participate in Quality Control exercises such as EQAP (External Quality Assurance Programme).

c) **Staff training:** All the staff involved in ELISA kit evaluation should undergo hands on training and competency testing on the following at suitable malaria labs before initiation of MDTL activity:

- Preparation and characterization of evaluation panel for the respective ELISA kit.
- Management of malaria ELISA kits received for performance evaluation (Verification/Storage/Unpacking etc).
- Perform tests, interpretation and documentation of results and reporting.
- Data management and safety and confidentiality

3. Reference sample panel and sample panel characterization:

To evaluate the performance of ELISA kit a well characterised malaria stage specific antigen/species specific antibody ELISA evaluation sample panel is required. In the absence of WHO Pre-Qualified/US FDA/ ATAGI Australia/ PMDA Japan approved malaria ELISA assay, it is recommended that performance evaluation of ELISA assays be carried out on a rigorously well characterized panel of positive and negative samples.

WHO Pre-Qualified/US FDA/ ATAGI Australia/ PMDA Japan approved ELISA kits may be used as the reference assay as and when such kits become available.

A statistically significant number of blood/plasma/sera samples should be used. All samples should be further characterized.

A. <u>Sample panel for malaria antigen ELISA evaluation:</u>

<u>Positive samples:</u> These are samples positive by RDT/microscopy **AND** PCR (standardized Snounou protocol/WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved) **AND** characterized further. These samples should be additionally characterized for parasite load and analyte level on in-house calibrated equipment using microscopy and other relevant test results.

a. <u>Range of Parasitemia:</u> Panel members should have low (≤200 parasites per microliter) to high (≥2000 parasites per microliter) range of *Plasmodium falciparum, P. vivax* and/or other *Plasmodium* species, as obtained from microscopy and/or other relevant test results. Characterized panels **must** contain equal number of samples of both low and high parasitemia.
b. <u>Analyte Characterization by ELISA:</u> Consistent ELISA quantification results should be obtained in ≥3 runs of ELISA experiments performed for each of the three antigens (PfHRP2, LDH and aldolase), with the results obtained at the 200

p/µL and the 2,000 p/µL being consistent with each other as well (factor of roughly 10 between results). The limit of detection of PfHRP2 is 0.6-74 ng/ mL, PvLDH is 1.6-47.9 ng/ mL, PfLDH is 0.2-53.5 ng/mL, and Pf aldolase is 0-9.9 ng/mL.

****** If the pool of samples available for testing is sufficiently large in numbers, then the antigen concentration range at the 200 p/µL dilution should be restricted to 5-9.5 ng/mL for PfHRP2, 15-47.9 ng/mL for PvLDH, 10.8-53.5 ng/mL for PfLDH, and 1.7-15ng/mL for Pf aldolase.

For those kits which have other antigen as target analyte (for which limits of detection have not been established), characterization of samples for that analyte should be performed on calibrated equipment in ≥3 runs of ELISA, leading to their classification as low/high parasitemic AND strong/moderate/weakly positive samples, which will then be used for performance evaluation of the assay.

B. <u>Sample panel for malaria antibody ELISA evaluation:</u>

For those kits which have antibodies as target analytes (for which limits of detection have not been established), characterization of sample panel for that analyte should be performed on calibrated equipment(≥3 runs of ELISA), leading to their classification as strong, moderate and weakly reactive samples and negative samples, which will then be used for performance evaluation of the assay. Samples for antibody kit evaluation must have been collected from microscopy/RDT/PCR positive malaria cases at requisite time points (depending on the antibody target).

The above-mentioned activities should not be performed with spiked/contrived samples.

Equal representation of samples positive for Plasmodium (P.falciparum /P.vivax) species is preferred.

C. *Negative panel for antigen based ELISA should constitute malaria RDT/microscopy negative samples (confirmed by PCR) as described in point 3A. Negative panel for antibody based ELISA should be negative for the analyte when performed on calibrated equipment (≥3 runs of ELISA).*

The reference sample panel should be stored in appropriate storage conditions (depending on the sample type and planned storage duration), and the quality of the panel should be checked periodically (at least once a year) with appropriate tests (e.g.: reference test/parasite culture/enzymatic activity/other relevant test).

Malaria positive samples should be obtained from health facilities, including tertiary care centers and their linked hospitals, private clinics, field practice areas etc.

4. **Sample size**: Sample sizes of positive and negative samples of each species targeted by the kit against different values of sensitivity and specificity are provided in Table 1 and Table 2, with recommended composition. Sample sizes have been calculated assuming 95% level of significance and an absolute precision of 5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity. Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ \left(1 - S_p\right)}{d^2}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*

Table 1. Positive sample sizes (per species) and composition for different values of sensitivity claimed by the manufacturer for evaluation of Pf (single/combo) or Pv (single/combo) antigen/antibody ELISA

Sensitivity	*Sample size: Minimum number of positive samples#*	*Composition of positive sample panel*
99%	16 (rounded to 20)	*Equal number of high and low parasitemic samples in each category for antigen based ELISA.* *It is recommended to include 30% strongly reactive samples, and 35% each across moderately and weakly reactive (depending on ELISA OD cut-offs) in each category for both antigen and antibody-based ELISA.*
95%	73 (rounded to 80)	
90%	139 (rounded to 140)	
85%	196 (rounded to 200)	
80%	246 (rounded to 250)	
75%	289 (rounded to 290)	

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Negative sample sizes and composition for different values of specificity claimed by the manufacturer for evaluation of Pf (single/combo) or Pv (single/combo) ELISA

Specificity	*Sample size: Minimum number of negative samples #*	*Composition of negative samples*
99%	16 (rounded to 20)	Dengue NS1/IgM positive samples: 03 Chikungunya IgM positive samples:03 Serum reactive for RA factor – low positive and high positive:02 Serum reactive for TPHA/other specific test for syphilis:02 Healthy controls from endemic regions: 10
95%	73 (rounded to 80)	Dengue NS1/IgM positive samples: 10 Chikungunya IgM positive samples:10 Serum reactive for RA factor – low positive and high positive:10 Serum reactive for TPHA/other specific test for syphilis:10 Healthy controls from endemic regions: 40
90%	139 (rounded to 140)	Dengue NS1/IgM positive samples: 18 Chikungunya IgM positive samples:18 Serum reactive for RA factor – low positive and high positive:18 Serum reactive for TPHA/other specific test for syphilis:18 Healthy controls from endemic regions: 68
85%	196 (rounded to 200)	Dengue NS1/IgM positive samples: 25 Chikungunya IgM positive samples:25 Serum reactive for RA factor – low positive and high positive:25 Serum reactive for TPHA/other specific test for syphilis:25 Healthy controls from endemic regions: 100
80%	246 (rounded to 250)	Dengue NS1/IgM positive samples: 30 Chikungunya IgM positive samples:30 Serum reactive for RA factor – low positive and high positive:30 Serum reactive for TPHA/other specific test for syphilis:30 Healthy controls from endemic regions: 130

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

5. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (comprising low and high analyte/parasitemic samples) and 3 negative samples per species (if applicable) should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

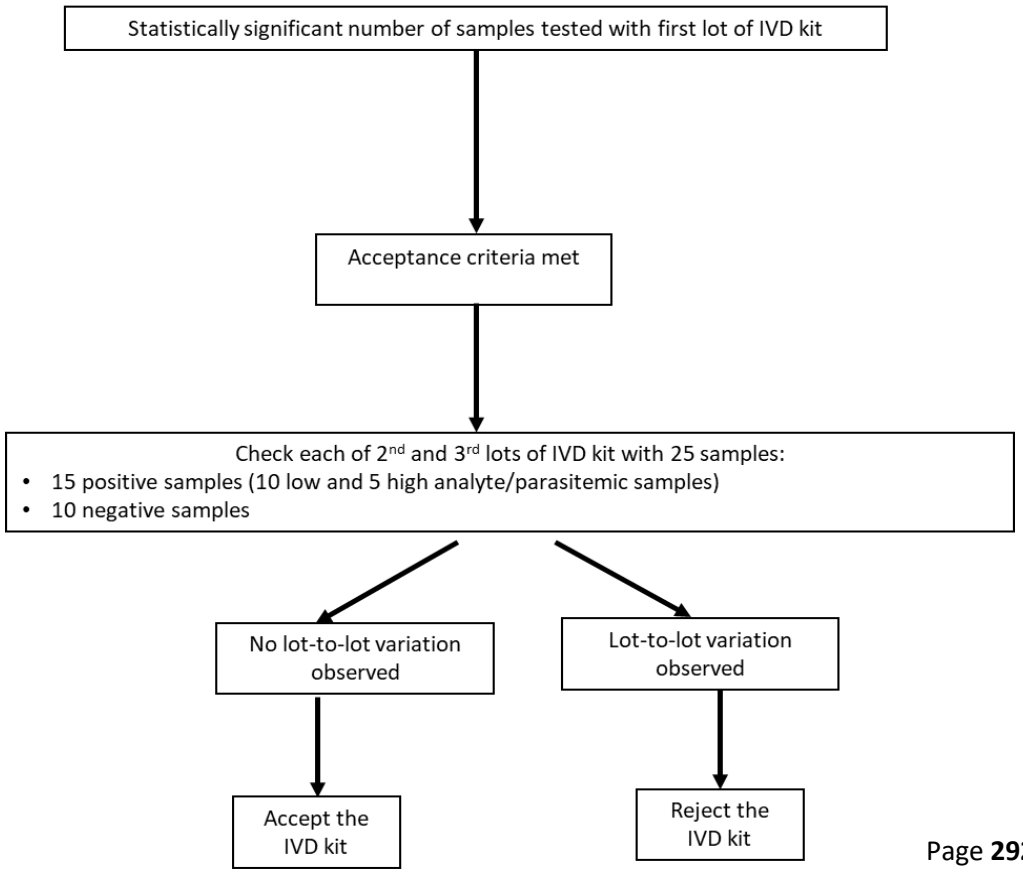
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low and 5 high analyte/parasitemic samples and 10 negative samples) per species (as applicable))
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low and 5 high analyte/parasitemic samples and 10 negative samples) per species (as applicable))
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Fig.1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (comprising low and high analyte/parasitemic samples) and 3 negative samples per species (as applicable) should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (comprising low and high analyte/parasitemic samples) and 3 negative samples per species (as applicable) should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (comprising low and high analyte/parasitemic samples) and 3 negative samples per species (as applicable) should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

6. Evaluation Methodology:

The index test should be tested on a rigorously well-characterized panel of samples from confirmed malaria positive and negative cases, which are also tested for the presence of malaria parasite using standardized Snounou protocol/microscopy/RDT or PCR Pre-Qualified by WHO or approved by US FDA/ATAGI Australia/PMDA Japan.

7. Interpretation of results:

Results should be interpreted as per the IFU of the reference assay.

8. Resolution of discrepant results:

True positive samples: These are well-characterized samples from confirmed malaria positive cases, which are also positive by the index test.
True negative samples: These are well-characterized samples from confirmed malaria negative cases, which are also negative by the index test.
False positive samples: These are well-characterized samples from confirmed malaria negative cases, which are positive by the index test.
False negative samples: These are well-characterized samples from confirmed malaria positive cases, which are negative by the index test.

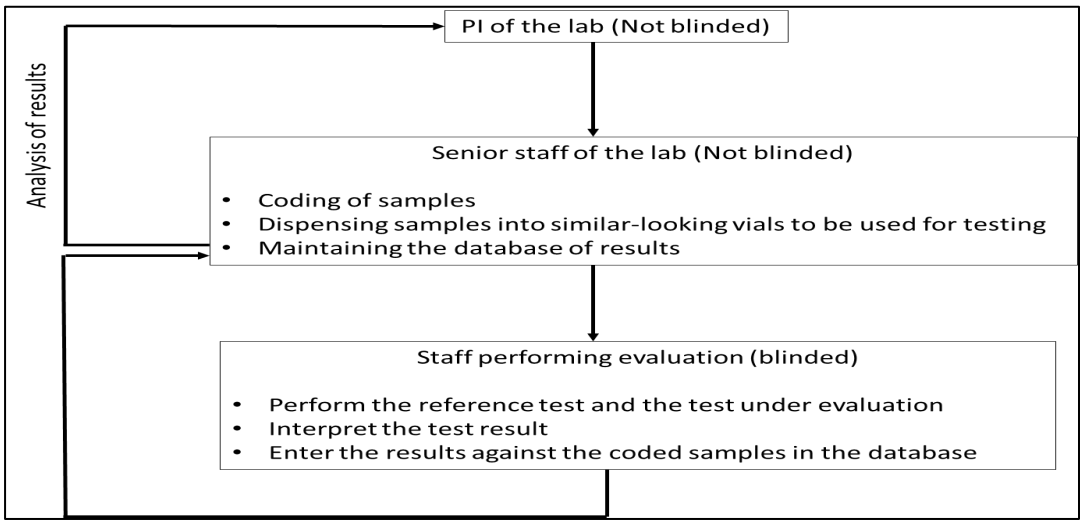
9. Acceptance Criteria:

Type of assay	Acceptance criteria	Minimum no. of samples needed to achieve at least the performance characteristics outlined in the acceptance criteria
Malaria antibody ELISA	Sensitivity: ≥90% Specificity: ≥95%	Minimum no. of Positive samples = 140 Minimum no. of Negative samples = 80
Pv ELISA	Sensitivity: ≥75% Specificity: ≥95%	Minimum no. of Positive samples = 290 Minimum no. of Negative samples = 80
Pf ELISA	Sensitivity: ≥90% Specificity: ≥95%	Minimum no. of Positive samples = 140 Minimum no. of Negative samples = 80
Cross-reactivity: Minimal		

10. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



11. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Snounou G, Viriyakosol S, Zhu XP, Jarra W, Pinheiro L, Do Rosario VE, et al. High sensitivity of detection of human malaria parasites by the use of nested polymerase chain reaction. Molecular and Biochemical Parasitology. 1993;61:315–20.
2. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w
3. World Health Organization. Methods manual for laboratory quality control testing of malaria rapid diagnostic test. Version Ten; 2023

VII. Performance evaluation report format

REPORT FORMAT

Name of the Laboratory

Name of the Institute, (with station)

Certificate of Analysis

File No.: _____

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number: Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: low/high parasitemia/simulated samples)	
	Negative samples (provide details, including cross reactivity panel/simulated samples)	

Results:

		Samples with confirmed disease status		
		Positive	Negative	Total
Name of malaria ELISA kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity
- Invalid test rate

- **Conclusions:**
  - Sensitivity, specificity
  - Invalid test rate
  - Performance: **Satisfactory / Not Satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting on ……. samples only, using kits provided by the manufacturer from the batch mentioned above. Results should not be extrapolated for any other sample type.)

Disclaimers

1. This validation process does not approve / disapprove the kit design

2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for …….. Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……

Performance evaluation protocol for Malaria real-time PCR kits

I. Background:

CDSCO/ICMR, New Delhi, have aimed to facilitate the evaluation and supply of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD kit performance.

II. Purpose:

To evaluate the performance characteristics of Malaria real-time PCR (RT-PCR) kits using irreversibly de-identified leftover archived clinical samples.

III. Requirements:

1. Instructions for use (IFU)
2. Supply of kits under evaluation (with batch no. and lot no. ; Manufacturing and Expiry and other required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
3. Evaluation sites/laboratories (With required equipment)
4. Reference test kits
5. Characterised Evaluation panel
6. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through

a) Laboratory accreditation: Accreditation for at least one of the Quality management systems (accreditation for Testing Lab / Calibration Lab (ISO/IEC: 17025), Medical Lab (ISO: 15189), PT provider (ISO/IEC: 17043) or CDSCO approved Reference laboratory.

b) It is recommended that malaria Medical Device Testing Labs (MDTLs) participate in Quality Control exercises such as EQAP (External Quality Assurance Programme).

c) **Staff training:** All the staff involved in IVD kit evaluation should undergo hands-on training and competency testing on the following at suitable malaria labs before initiation of MDTL activity:

- Preparation and characterization of evaluation panel for the respective IVD kit.
- Management of RDT kits (specific for *Plasmodium falciparum / Plasmodium vivax*) received for performance evaluation (Verification/Storage/Unpacking etc.).
- Perform tests interpretation and documentation of results, and reporting.
- Data management and safety and confidentiality.

3. Preparation of evaluation sample panel for Malaria

To evaluate the performance of malaria RT-PCR IVD kit, a well characterized species specific malaria whole genome panel is required. Hence, statistically significant number of whole blood samples should be used. *The panel should comprise positive and negative samples as described in section 8.*

The reference sample panel should be stored in appropriate storage conditions (depending on the sample type and planned storage duration), and the quality of the panel should be checked periodically (at least once a year) with appropriate tests (e.g.: reference test/parasite culture/enzymatic activity/other relevant test).

Malaria positive samples should be obtained from health facilities, including tertiary care centers and their linked hospitals, private clinics, field practice areas etc.

4. *DNA extraction*

DNA extraction should be performed using a standard protocol/kit as recommended by the manufacturer, or fully automated DNA extractor may be used (as per manufacturer's instruction and compatible reagent kits).

Note: If the manufacturer of the index test recommends a specific DNA extraction kit, it needs to be provided by the manufacturer, if the evaluation lab is unable to procure the same.

5. Real-time PCR system:

PCR should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Real-time closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.

6. Internal Control/Extraction Control:

The index test must have an internal control (housekeeping gene), with or without an extraction control.

7. Reference assay:

Two WHO Pre-Qualified/ US FDA/ ATAGI Australia/ PMDA Japan-approved malaria RT-PCR assays (or one approved assay from the aforementioned regulatory authorities and

standardized Snounou protocol) should be used as reference assays for the characterization of samples, with 100% agreement between their results.

All positive samples should be confirmed positive by the reference assay(s).

All negative samples should be confirmed negative by the reference assay(s).

8. Sample size and sample panel composition for performance evaluation:

Sample sizes of positive and negative samples of each species targeted by the kit against different values of sensitivity and specificity are provided in Table 1 and Table 2, with recommended composition. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate of 5%. Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity. Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Table 1. Positive sample sizes (per species) and composition for different values of sensitivity claimed by the manufacturer for evaluation of malaria real time PCR kit

Sensitivity	*Sample size: Minimum number of positive samples per species#*	*Composition of positive samples*
99%	16 (rounded to 20)	Equal number of high and low parasitemic samples

95%	77 (rounded to 80)	Equal number of high and low parasitemic samples
90%	146 (rounded to 150)	Equal number of high and low parasitemic samples
85%	207 (rounded to 210)	Equal number of high and low parasitemic samples
80%	259 (rounded to 260)	Equal number of high and low parasitemic samples
75%	304 (rounded to 310)	Equal number of high and low parasitemic samples

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Table 2. Negative sample sizes and composition for different values of specificity claimed by the manufacturer for evaluation of malaria real time PCR kit

Specificity	*Sample size: Minimum number of negative samples per species#*	*Composition of negative samples*
99%	16 (rounded to 20)	Dengue NS1/IgM positive samples: 03 Chikungunya IgM positive samples:03 Serum reactive for RA factor – low positive and high positive:02 Serum reactive for TPHA/other specific test for syphilis:02 Healthy controls from endemic regions: 10
95%	77 (rounded to 80)	Dengue NS1/IgM positive samples: 10 Chikungunya IgM positive samples:10 Serum reactive for RA factor – low positive and high positive:10 Serum reactive for TPHA/other specific test for syphilis:10 Healthy controls from endemic regions: 40
90%	146 (rounded to 150)	Dengue NS1/IgM positive samples: 18 Chikungunya IgM positive samples:18 Serum reactive for RA factor – low positive and high positive:18 Serum reactive for TPHA/other specific test for syphilis:18 Healthy controls from endemic regions: 78
85%	207 (rounded to 210)	Dengue NS1/IgM positive samples: 26 Chikungunya IgM positive samples:26 Serum reactive for RA factor – low positive and high positive:26 Serum reactive for TPHA/other specific test for syphilis:26 Healthy controls from endemic regions: 106
80%	259 (rounded to 260)	Dengue NS1/IgM positive samples: 35 Chikungunya IgM positive samples:35 Serum reactive for RA factor – low positive and high positive:30 Serum reactive for TPHA/other specific test for syphilis:30 Healthy controls from endemic regions: 130

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Sample panel composition:

A. ***Positive samples:*** *RDT/microscopy positive malaria samples should be obtained from health facilities and confirmed* using two USFDA/ATAGI Australia/PMDA Japan approved/WHO Pre-Qualified PCR Kits (or one approved kit from any of the aforementioned authorities and standardized Snounou protocol). Once the positive samples are well-characterized with these two PCR assays (100% agreement between results), they should be classified as per their parasite load using microscopy and/or other relevant test format.
Range of Parasitemia: Panel members should have a low (≤200 parasites per microliter) to high (≥2000 parasites per microliter) range of *Plasmodium falciparum, P. vivax*, as obtained from microscopy and/or other relevant test results. Characterized panels **must** contain equal number of samples of both low and high parasitemia.

Note for additional characterization (not mandatory):

*If ELISA is used for characterization of samples in addition to the above-mentioned mandatory tests (RDT/microscopy **AND** PCR), consistent ELISA quantification results should be obtained in ≥3 runs of ELISA experiments performed for each of the three common antigens (PfHRP2, LDH and aldolase), with the results obtained at the 200 p/µL and the 2,000 p/µL being consistent with each other as well (factor of roughly 10 between results). The limit of detection of PfHRP2 is 0.6-74 ng/ mL, PvLDH is 1.6-47.9 ng/ mL, PfLDH is 0.2-53.5 ng/mL, and Pf aldolase is 0-9.9 ng/mL.*

******* If the pool of samples available for testing is sufficiently large in numbers, then the antigen concentration range at the 200 p/µL dilution should be restricted to 5-9.5 ng/mL for PfHRP2, 15-47.9 ng/mL for PvLDH, 10.8-53.5 ng/mL for PfLDH, and 1.7-15ng/mL for Pf aldolase.*

If ELISA for other antigen is carried out for further characterization (for which limits of detection have not been established), testing should be carried out on calibrated equipment in ≥3 runs of ELISA, leading to their classification as low and high parasitemic samples or strong/moderate/weakly reactive samples.

Negative samples: These are samples negative by RDT/microscopy **AND** two USFDA/ATAGI Australia/PMDA Japan approved/WHO Pre-Qualified PCR Kits (or one approved kit from any of the aforementioned authorities and standardized Snounou protocol).

The above-mentioned activities should not be performed with spiked/contrived samples.

If the kit does not differentiate between Pf and Pv, performance characteristics may be evaluated for the kit as a whole with statistically significant sample size according to different performance metrics outlined in Tables 1 and 2, with equal representation of samples positive for P.falciparum and /P.vivax.

It is desirable (but not mandatory) to evaluate the kit using some positive samples from mixed Pf/Pv infection cases.

9. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (comprising low and high parasitemic samples) and 3 negative samples per target pathogen should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

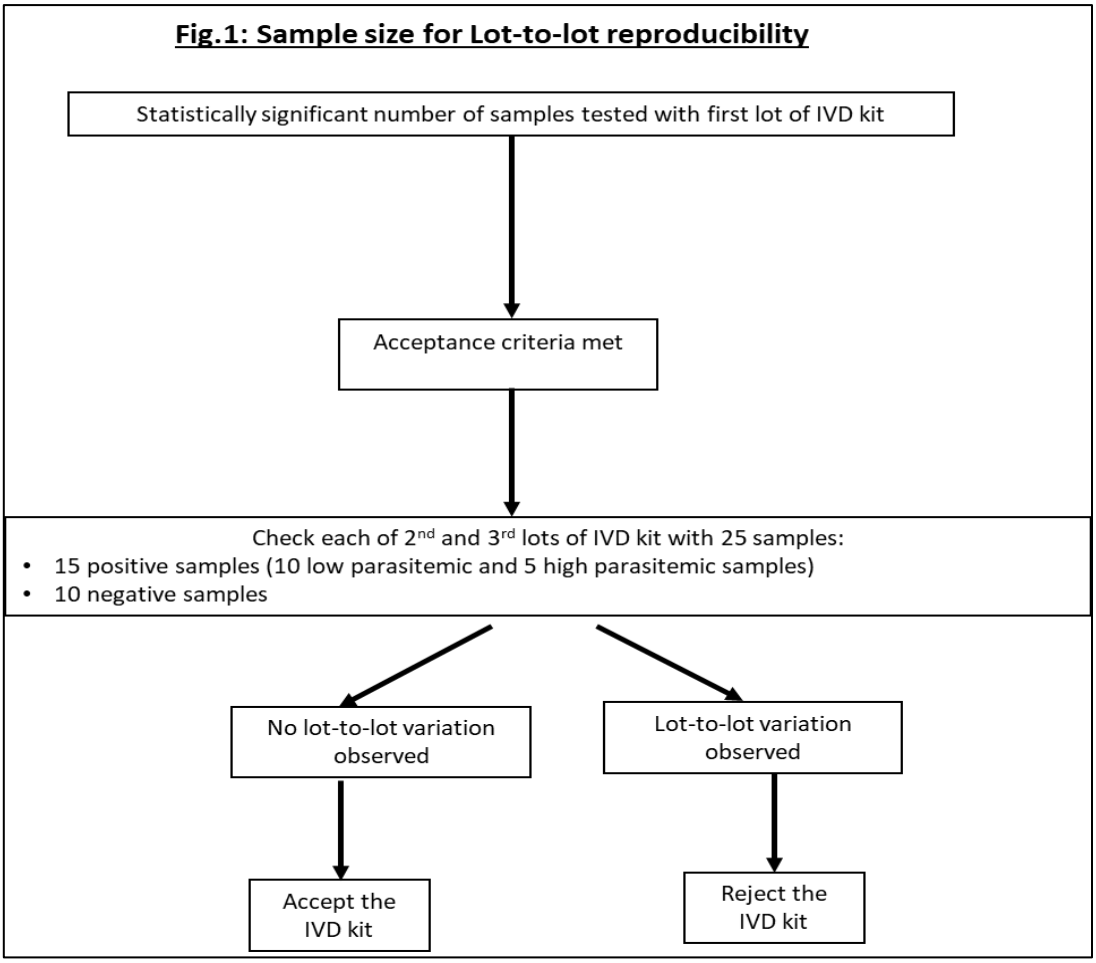
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low parasitemic AND 5 high parasitemic samples, and 10 negative samples) per target pathogen.
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low parasitemic AND 5 high parasitemic samples, and 10 negative samples) per target pathogen.
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Fig.1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (comprising low and high parasitemic samples) and 3 negative samples per target pathogen should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (comprising low and high parasitemic samples) and 3 negative samples per target pathogen should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (comprising low and high parasitemic samples) and 3 negative samples per target pathogen should be tested 5 times in independent runs.
   Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report Ct standard deviation across repeats (≤1cycle deviation recommended).

10. Testing Methodology:

The reference assay and the index test should be run on the sample panel in parallel.

11. Interpretation of results:

Results should be interpreted as per the IFU of the reference assay and the index test.

12. Resolution of discrepant results:

True positive samples: These are samples positive by both reference assay and index test.
True negative samples: These are samples negative by both reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

13. Acceptance Criteria:

Target Plasmodium species	**Acceptance criteria**	**Minimum no. of samples needed to achieve at least the performance characteristics outlined in the acceptance criteria**
Pf PCR	Sensitivity ≥98% Specificity ≥98% Limit of detection: 1 parasite/µl Invalid test rate: ≤5%	Minimum no. of Positive samples = 80 Minimum no. of Negative samples = 80
Pv PCR	Sensitivity ≥95% Specificity ≥98% Limit of detection: 1-2 parasites/µl Invalid test rate: ≤5%	Minimum no. of Positive samples = 80 Minimum no. of Negative samples = 80
Multiplex PCR - Pf & Pv	For Pf: • Sensitivity: ≥98% • Specificity: ≥98% • Absolute precision 5% • 95% CI • Invalid test rate ≤5% • Limit of detection: 1 parasite/µl For Pv: • Sensitivity: ≥95% • Specificity: ≥98% • Absolute precision 5% • 95% CI • Invalid test rate ≤5%	<u>For Pf:</u> Minimum no. of Positive samples = 80 Minimum no. of Negative samples = 80 <u>For Pv:</u> Minimum no. of Positive samples = 80

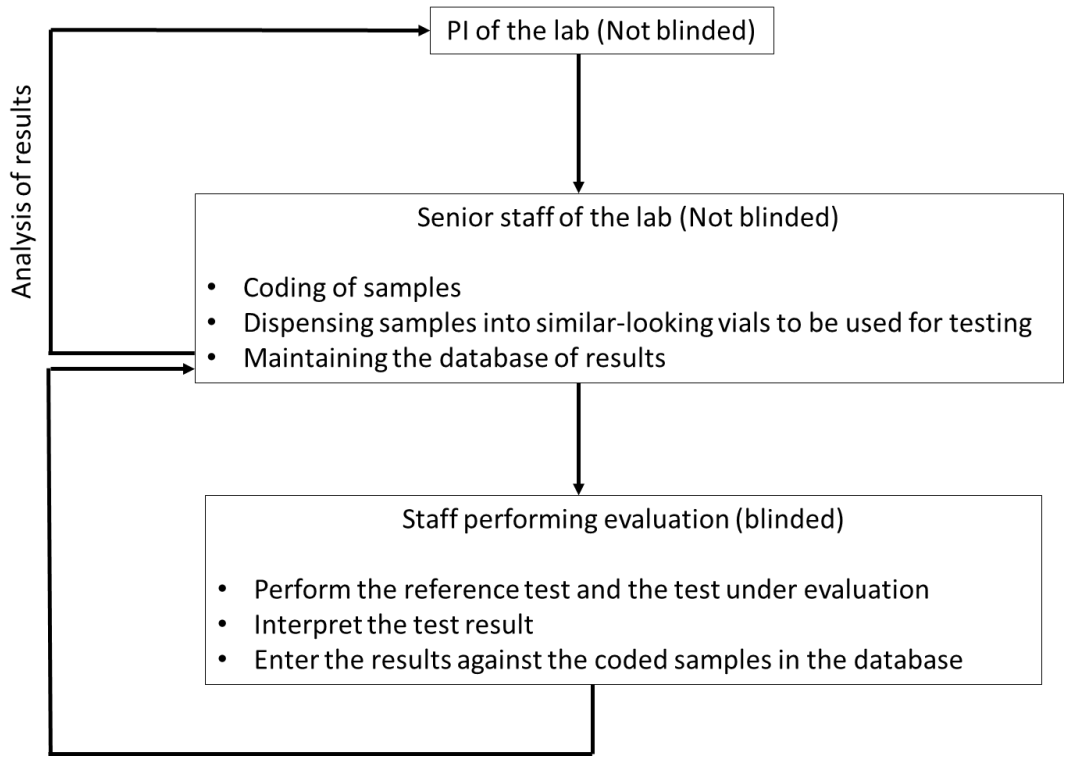
	• Limit of detection: 1-2 parasites/µl	Minimum no. of Negative samples = 80

Cross-reactivity: Minimal
Invalid test rate: ≤5%

14. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



15. Publication Rights

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Snounou G, Viriyakosol S, Zhu XP, Jarra W, Pinheiro L, Do Rosario VE, et al. High sensitivity of detection of human malaria parasites by the use of nested polymerase chain reaction. Molecular and Biochemical Parasitology. 1993;61:315–20.
2. Ramírez AM, Tang THT, Suárez ML, Fernández AÁ, García CM, Hisam S, Rubio JM. Assessment of Commercial Real-Time PCR Assays for Detection of Malaria Infection in a Non-Endemic Setting. Am J Trop Med Hyg. 2021 Oct 12;105(6):1732-1737. doi: 10.4269/ajtmh.21-0406. PMID: 34662870; PMCID: PMC8641344.
3. Bouzayene, A., Zaffaroullah, R., Bailly, J. *et al.* Evaluation of two commercial kits and two laboratory-developed qPCR assays compared to LAMP for molecular diagnosis of malaria. *Malar J* **21**, 204 (2022). https://doi.org/10.1186/s12936-022-04219-1
4. Aschar M, Sanchez MCA, Costa-Nascimento MJ, Farinas MLRN, Hristov AD, Lima GFMC, Inoue J, Levi JE, Di Santi SM. Ultrasensitive molecular tests for *Plasmodium* detection: applicability in control and elimination programs and reference laboratories. Rev Panam Salud Publica. 2022 Mar 28;46:e11. doi: 10.26633/RPSP.2022.11. PMID: 35355692; PMCID: PMC8959250.
5. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w
6. World Health Organization. Methods manual for laboratory quality control testing of malaria rapid diagnostic test. Version Ten; 2023

VII. Performance evaluation report format

REPORT FORMAT

Name of the Laboratory

Name of the Institute, (with station)

Certificate of Analysis

File No.: _____

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number: Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: low/high parasitemia/simulated samples)	
	Negative samples (provide details, including cross reactivity panel/simulated samples)	

Results:

		Reference assay …….…… (name)		
		Positive	Negative	Total
Name of malaria real time PCR kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity
- Invalid test rate

- **Conclusions:**
  - Performance: **Satisfactory / Not Satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting on ……. samples only, using kits provided by the manufacturer from the batch mentioned above. Results should not be extrapolated for any other sample type.)

Disclaimers

1. This validation process does not approve / disapprove the kit design

2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for …….. Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……

Field evaluation protocol for combo Malaria Rapid Diagnostic Test (RDT) kits (detecting *P vivax* and *P falciparum*)

I. Background:

CDSCO/ICMR, New Delhi, have aimed to facilitate the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD kit performance.

II. Purpose:

To evaluate the performance characteristics of Malaria RDT kits (detecting *P. vivax* and/or *P. falciparum*) in the diagnosis of Malaria parasite infection in individuals with unknown disease status.

III. Requirements:

1. Supply of kits under evaluation (with batch no. and lot no. Manufacturing and Expiry dates other required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Laboratory supplies

IV. Ethical approval:

The study will be initiated after approval from the institutional human ethics committee.

V. Procedure:

1. **Study design/type**: Cross-sectional study
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through
   A. Laboratory accreditation: Accreditation for at least one of the Quality management systems (accreditation for Testing Lab / Calibration Lab (ISO/IEC: 17025), Medical Lab (ISO: 15189), PT provider (ISO/IEC: 17043) or CDSCO approved Reference laboratory.

   It is recommended that malaria Medical Device Testing Labs (MDTLs) participate in Quality Control exercises such as EQAP (External Quality Assurance Programme).

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on the following at suitable malaria labs before initiation of MDTL activity:

   - Preparation and characterization of evaluation panel for the respective IVD kit.
   - Management of RDT kits (specific for *Plasmodium falciparum / Plasmodium vivax*) received for performance evaluation (Verification/Storage/Unpacking etc.).

- Perform tests interpretation and documentation of results, and reporting.
- Data management and safety and confidentiality.

3. Sample size for performance evaluation:

Sample sizes of positive and negative samples against different values of sensitivity and specificity are provided in Tables 1 and 2. Sample sizes have been calculated assuming 95% level of significance, an absolute precision of 5%, and invalid test rate 5%. It is further assumed that at least 5% of the individuals attending the health care facilities for acute febrile illness and suspected for Malaria will be positive for Malaria (*P. vivax* and *P. falciparum*). Appropriate sample size has to be chosen from the tables according to the values of sensitivity and specificity being claimed by the manufacturer. If a claimed sensitivity/specificity is not present in the table, the manufacturer needs to consider the sample size associated with the largest sensitivity/specificity provided in the table that is smaller to the claimed value (that is, as per the next smaller value of the sensitivity/ specificity available in the table). For example, if a manufacturer claims a sensitivity of 93%, they are required to use a sample size mentioned against 90% sensitivity. Similarly, a claim of 87% specificity would require usage of the sample size outlined for 85% specificity. Sample sizes are calculated using the following formulae and assumption of 5% for prevalence of the disease:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR}) \text{ x P}}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR}) \text{ x P}}$$

- *n (se) is the minimum number of individuals to be enrolled to obtain the requisite number of positive samples.*
- *n (sp) is the minimum number of individuals to be enrolled to obtain the requisite number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*
- *P is prevalence of the disease*

Sample size has to be calculated based on both the sensitivity and the specificity. The final sample size will be the maximum of the two. For example, at 95% sensitivity and 95% specificity, the sample size required will be 1600 (maximum of 1600 and 84). Please note that since the prevalence is low, the final sample size is generally expected to be governed by the assumed sensitivity.

Table 1. Sample sizes for different values of species-specific sensitivity being claimed

Sensitivity	Minimum no. of positive samples required (rounded figure) #	Minimum number of individuals to be enrolled in the study to obtain requisite number of positive samples
99%	20	400
95%	80	1600
90%	150	3000
85%	210	4200
80%	260	5200
75%	305	6100

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Samples will be collected from individuals attending the health care facilities *(tertiary care centers and their linked hospitals, private clinics, field practice areas etc.)* for acute febrile illness in highly endemic areas.

The disease status of these cases will be unknown.

Table 2. Sample sizes for different values of species-specific specificity being claimed

Specificity	No. of negative samples required (rounded figure)	Minimum number of individuals to be enrolled to obtain requisite number of negative samples
99%	20	21
95%	80	84
90%	150	158
85%	210	221
80%	260	274
75%	305	321

#It is recommended to calculate the sample size as per manufacturer's claims of sensitivity and specificity; however, a higher sample size is suggested to ensure adequate power of the study in case the kit falls short of claimed performance characteristics.

Samples will be collected from individuals attending the health care facilities *(tertiary care centers and their linked hospitals, private clinics, field practice areas etc.)* for acute febrile illness in highly endemic areas.

Since a large number of febrile cases have to be enrolled to obtain the requisite number of malaria positive samples, enrolling the number of cases mentioned in Table 1 will be sufficient to obtain the requisite number of negative samples.

4. Inclusion criteria:

Individuals with the following clinical features may be enrolled in the study

Fever and any 2 of the following:

- Chills, sweating, headache, tiredness, nausea and vomiting, jaundice, splenomegaly

5. Exclusion criteria

- Individuals not satisfying inclusion criteria
- Individuals with already known positive history for other pathogens

6. Reference assay:

WHO Pre-Qualified/ US FDA/ ATAGI Australia/ PMDA Japan approved Malaria PCR assay/ standardized Snounou protocol should be used as reference assay.

7. Study implementation:

The patients displaying Malaria like illness will be recruited into the study and five ml of whole blood will be collected in EDTA tubes. The whole blood sample will be subjected to the reference and the index test.

The disease status of the enrolled cases will be unknown.

8. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

9. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

10. Positive samples:

Samples positive by the reference assay will be considered as true positive samples.

11. Negative samples:

Samples negative by the reference assay will be considered as true negative samples.

False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

12. Cross reactivity:

The RDT kit should have been evaluated against the following cross reactivity panel during the analytical performance evaluation:

- *Dengue NS1 positive samples (n=10 samples)*
- *Chikungunya PCR positive samples (n=10 samples)*
- *Healthy controls from endemic regions (n= 40 samples)*
- *Serum reactive for RA factor – low positive and high positive (n=15 samples)*
- *Serum reactive for TPHA/other specific test for syphilis (n= 10 samples)*

13. Statistical analysis:

Sensitivity and specificity will be calculated.

Interim analysis of data shall be conducted on completing evaluation of 25%, 50% and 75% of samples. If, at any point, the performance of the assay is found to be not satisfactory, the assay shall not be evaluated further. Evaluation fee shall be charged accordingly.

14. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target analyte using the kit under evaluation.

3 positive samples (comprising low parasitemic AND high parasitemic samples) and 3 negative samples per species should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

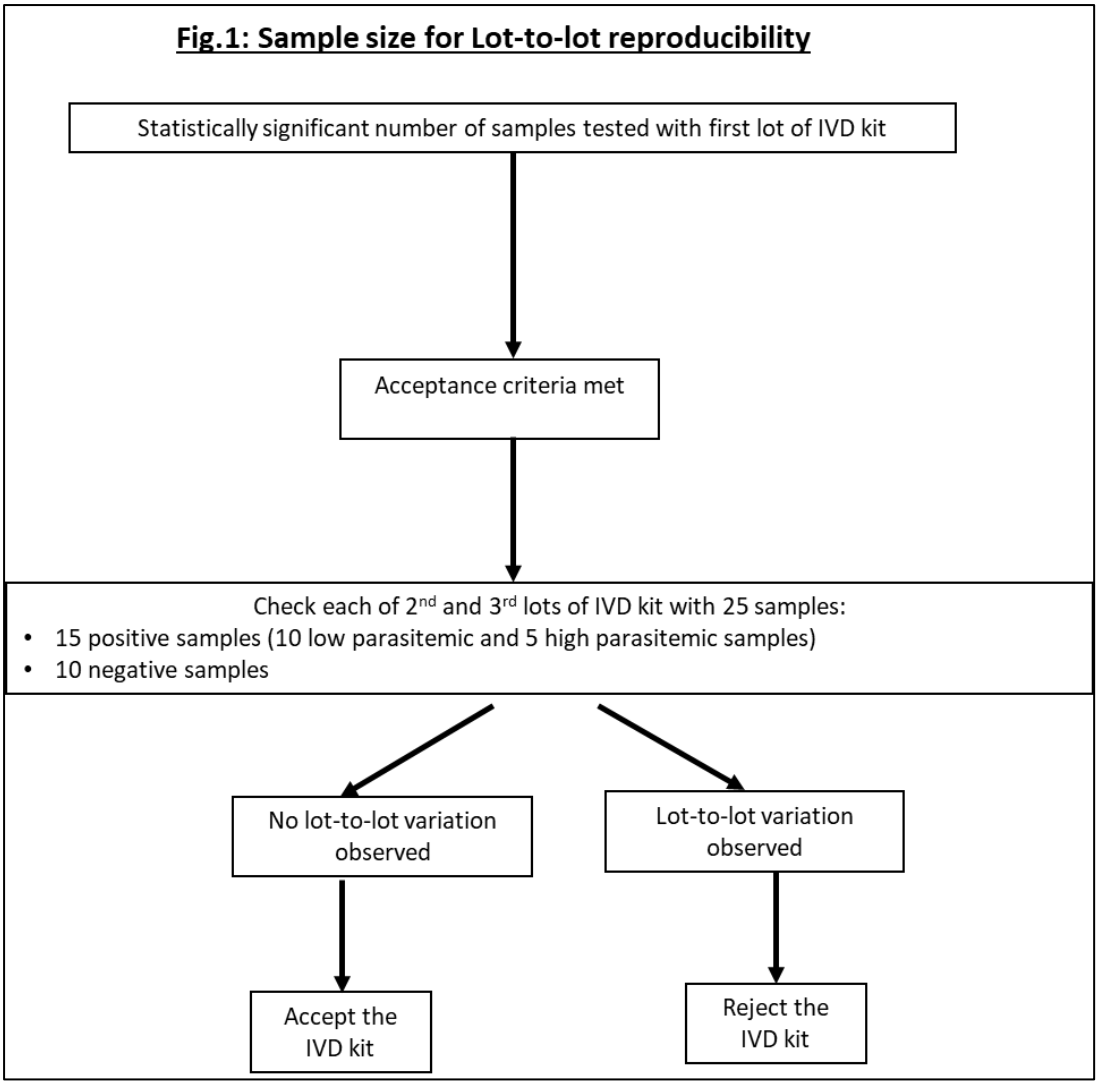
B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low parasitemic AND 5 high parasitemic samples) per species, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low parasitemic AND 5 high parasitemic samples) per species, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (comprising low parasitemic AND high parasitemic samples) and 3 negative samples per species should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (comprising low parasitemic AND high parasitemic samples) and 3 negative samples per species should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3

positive samples (comprising low parasitemic AND high parasitemic samples) and 3 negative samples per species should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

14. Resolution of discrepant results:

True positive samples: These are samples positive by both reference assay and index test.
True negative samples: These are samples negative by both reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

15. Blinding of laboratory staff

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the results of the reference assay. The PI of the evaluation exercise should remain unblinded, i.e., privy to the results of the reference test. Another senior laboratory staff selected by the PI may remain unblinded for overseeing the activity and maintaining the database of results. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data.

16. Acceptance criteria:

Sensitivity: ≥75% for *P. vivax* and ≥95% for *P. falciparum*
Specificity: ≥90% for *P. vivax* and ≥95% for *P. falciparum*
Cross-reactivity: Minimal
Invalid test rate: ≤5%

To achieve at least the performance characteristics outlined in the acceptance criteria for *P vivax*, ≥6100 individuals satisfying the case definition need to be enrolled to obtain the requisite number of positive samples. This sample size is sufficient for requisite number of negative samples.
To achieve at least the performance characteristics outlined in the acceptance criteria for *P falciparum*, ≥1600 individuals satisfying the case definition need to be enrolled to obtain the requisite number of positive samples. This sample size is sufficient for requisite number of negative samples.

Recruitment should be terminated once the desired number of positive cases is enrolled and tested.

17. Publication Rights

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Ministry of Health and Family Welfare. Guidelines for Bivalent RDT. Available at: guidelines-for-bivalent-rdt.pdf (mohfw.gov.in)
2. World Health Organization. Malaria Rapid Diagnostic Test Performance - Results of WHO product testing of malaria RDTs: round 8 (2016–2018): Available at: https://iris.who.int/bitstream/handle/10665/276190/9789241514965-eng.pdf?sequence=1
3. Snounou G, Viriyakosol S, Zhu XP, Jarra W, Pinheiro L, Do Rosario VE, et al. High sensitivity of detection of human malaria parasites by the use of nested polymerase chain reaction. Molecular and Biochemical Parasitology. 1993;61:315–20.
4. Integrated Disease Surveillance Project Training Manual For State & District Surveillance Officers - Case Definitions Of Diseases & Syndromes Under Surveillance (Module-5). Available at: https://idsp.mohfw.gov.in/WriteReadData/OldSite/2WkDSOSept08/Resources_files/DistrictSurvMan/Module5.pdf [Accessed on 25th June 2024]
5. CDC. National Notifiable Diseases Surveillance System (NNDSS). Malaria (Plasmodium spp.) 2014 Case Definition. Available at: https://ndc.services.cdc.gov/case-definitions/malaria-2014/ [Accessed on 28th June, 2024]
6. Kannambath R, Rajkumari N, Sivaradjy M. Prevalence of malaria: A 7-year trend analysis from a tertiary care center, Puducherry. Trop Parasitol. 2023 Jan-Jun;13(1):28-33. doi: 10.4103/tp.tp_41_22. Epub 2023 May 19. PMID: 37415756; PMCID: PMC10321582.
7. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

REPORT FORMAT

Name of the Laboratory

Name of the Institute, (with station)

Certificate of Analysis

File No.: _____

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number: Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples: Not applicable, may categorize cases as per duration of illness	
	Negative samples (may categorize as per duration of illness, must include cross reactivity panel)	

Results:

		Reference assay ……. (name)		
		Positive	Negative	Total
Name of index malaria RDT	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity
- Invalid test rate

- Conclusions:
  - Performance: Satisfactory / Not Satisfactory

(Sensitivity and specificity have been assessed in field/controlled lab setting on……. samples only, using kits provided by the manufacturer from the batch mentioned above. Results should not be extrapolated for any other sample type.)

Disclaimers

1. This validation process does not approve / disapprove the kit design

2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for …….. Kit (Lot No……) manufactured by …… (Supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……

Information on Operational and Test Performance Characteristics Required from Manufacturers for Malaria IVD

The manufacturer should provide the following details about the IVD:

1. Instructions for Use
2. Scope of the IVD: to diagnose Malaria (Pf and/or Pv)
3. Intended Use Statement
4. Principle of the assay
5. Intended testing population (cases of acute febrile illness/suspected cases of Malaria)
6. Intended user(laboratory professional and/or health care worker at point-of-care)
7. Detailed test protocol
8. Lot/batch No.
9. Date of manufacture
10. Date of Expiry
11. Information on operational Characteristics
    - i. Configuration of the kit/device
    - ii. Requirement of any additional equipment, device
    - iii. Requirement of any additional reagents
    - iv. Operation conditions
    - v. Storage and stability before and after opening
    - vi. Internal control provided or not
    - vii. Quality control and batch testing data
    - viii. Biosafety aspects- waste disposal requirements
12. Information on Test Performance Characteristics
    - i. Type of sample-serum/plasma/whole blood/other specimen (specify)
    - ii. Volume of sample
    - iii. Sample rejection criteria (if any)
    - iv. Any additional sample processing required
    - v. Any additional device/consumable like sample transfer device, pipette, tube, etc required
    - vi. Name of analyte to be detected
    - vii. Pathogens targeted by the kit

viii. Time taken for testing

ix. Time for result reading and interpretation

x. Manual or automated (equipment)reading

xi. Limit of detection/Limit of Quantification and range of detection

xii. Diagnostic sensitivity

xiii. Diagnostic specificity

xiv. Stability and reproducibility (including data)

xv. Training required for testing (if any)

xvi. If yes, duration

xvii. Details of Cut-off and /or Equivocal Zone for interpretation of test

xviii. Details of cross reactivity, if any

xix. Interpretation of invalid and indeterminate results to be provided

xx. It is recommended to provide data demonstrating accuracy and precision

xxi. Limit of detection

*****Please mention “Not applicable” against sections not pertaining to the kit.

NIPAH VIRUS REAL TIME PCR KIT

List of Contributors:

A. Working Group:

1. Dr. Anita Shete-Aich, Scientist-E, ICMR National Institute of Virology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
2. Dr. Labanya Mukhopadhyay, Scientist-C, Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India
3. Ms. Krittika Bhattacharyya, Statistical Officer (Planning), Directorate of Economics and Statistics, Government of National Capital Territory of Delhi

B. Review Committee:

1. Dr. Vasanthapuram Ravi, Former Dean Research and Head of Neurovirology, National Institute of Mental Health and Neuro-Sciences, Bengaluru, Karnataka
2. Dr. Pragya D Yadav, Scientist-F, ICMR National Institute of Virology and Director-in-Charge, National Institute of One Health, Department of Health Research, Ministry of Health and Family Welfare, Government of India
3. Mr. Pramod Meshram, Deputy Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
4. Dr. Sella Senthil, Assistant Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
5. Dr. Md Omair Anwar, Drugs Inspector (Medical Devices), IVD Division, Central Drugs Standard Control Organization, Government of India
6. Dr. Nivedita Gupta, Scientist-G and Head of the Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

Performance evaluation protocol for Nipah virus real-time PCR kit

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding in-vitro diagnostic kit (IVD) performance.

This recommendation focuses on the laboratory performance evaluation of Nipah virus real time PCR kit. All clinical samples tested in the study should be evaluated in accordance with the candidate test's instructions for use.

II. Purpose:

To evaluate the performance characteristics of Nipah virus real-time PCR kits in the diagnosis of Nipah virus infection/ disease using irreversibly de-identified leftover archived/ spiked clinical samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using spiked/clinical samples (human specimens).
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should be well-equipped and establish their proficiency through ALL of the following:

A. Availability of BSL-4 facility for handling of Nipah virus positive specimens

B. Accreditation for at least one Quality management system for at least one respiratory viral pathogen molecular testing (accreditation for Testing Lab / Calibration Lab as per ISO/IEC 17025, Medical Lab as per ISO 15189, PT provider as per ISO/IEC 17043), or CDSCO approved Reference laboratory.

C. Staff training: All the staff involved in Nipah virus IVD evaluation should undergo hands on training and competency testing on following

- BSL-4 practices
- Nipah virus culture and handling
- Preparation & characterization of reference sample panel
- Handling of Nipah virus RT-PCR kits received for performance evaluation (Verification/Storage/Unpacking etc).
- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Nipah virus RNA evaluation panel

This is a zoonotic disease, and well characterised Nipah virus positive human samples is a critical requirement for evaluation of RT-PCR IVD kits. A statistically significant number of clinical samples should be used for the evaluation.

4. RNA extraction

RNA extraction should be performed as per reference test and index test Instructions for Use (IFU).

If the manufacturer of the index test recommends a specific RNA extraction kit/system, the same needs to be provided by the manufacturer if the evaluating lab is unable to procure the same/it is not available within the lab's IVD evaluation scope.

5. Real-Time PCR System

PCR should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Real-time closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.

6. Internal control/Extraction control

Assays must have an internal control (housekeeping gene), with or without an extraction control (RNA added before extraction to a sample).

7. Reference assay:

The Nipah virus Real Time PCR Assay developed by ICMR-NIV Pune, or a WHO Pre-Qualified/ US FDA/ ATAGI Australia/ PMDA Japan approved real time PCR assay should be used as the Reference Standard.

All positive samples should be confirmed positive by the reference assay.

All negative samples should be confirmed negative by the reference assay.

8. Sample size for performance evaluation: Sample size is calculated assuming 95% sensitivity and specificity of the index test, 95% confidence level, absolute precision of 5% and ≤5% invalid test rate. A minimum of 77 (rounded to 80) positive clinical samples and a minimum of 77 (rounded to 80) negative clinical samples are required. Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

Nipah virus is detectable from throat swab, urine, CSF. The assay should be validated with positive clinical/spiked samples, and negative samples for all the formats claimed by the manufacturer. However, if a particular sample matrix is used to evaluate the assay (as opposed to all the sample types claimed by the manufacturer), the performance evaluation report should clearly mention the performance characteristics of the assay against the sample type used for validation. There should be no ambiguity about the sample type used for assay validation.

9. Sample panel composition:

A. Human samples

A.1 Positive samples (Minimum n=80 for each sample type): Clinical/ Spiked samples positive by the reference real-time PCR assay

A.1.1 Strong positive (Ct value <25) = 24 samples

A.1.2. Moderate positive (Ct value between 25-30) = 28 samples

A.1.3 Weak positive (Ct value >30 to 34) = 28 samples

The sample type should be as per the index test IFU. If an assay claims to detect Nipah virus RNA in several sample types, attempt should be made to use 80 positive samples across each sample type, or at least the sample types available with the evaluating lab. This relaxation is provided since clinical samples are scarce and obtained only during outbreaks occurring every few years in India, which necessitates using spiked clinical samples. The latter is difficult since Nipah virus is a BSL-4 level pathogen and its handling requires sophisticated laboratory setup and trained manpower.

In case the requisite number of specimens for a particular sample type are not available and a smaller number of samples are used for performance evaluation (i.e., sample size calculated assuming higher performance characteristics), it is necessary to ensure that the study has adequate power for acceptance of the evaluation results in case the assay falls short of the assumed performance characteristics.

Note:

If clinical samples positive for Nipah virus are not available, tissue culture fluid (Heat-inactivated) from reference laboratories can be used, spiked in serum/urine/Throat swab samples to obtain the panel with Ct value <25, 25-30 and >35 and tested by the reference assay, and the positive samples can be used for evaluation.

Confirmed negative samples would be used for spiking with Nipah virus.isolate.

A.2 Negative samples (number of samples will depend on sample type): All negative samples should be negative by reference real-time PCR assay. Distribution of the negative samples should be as follows

Categories of samples as per the sample type	Sample type		
	NP/TS (Minimum n= 80)	Serum (Minimum n= 80)	Urine (Minimum n=80)
A.2.1 Samples from cases having similar illness/ spiked samples which are RT-PCR positive for common pathogens but negative for Nipah virus	Samples from individuals presenting with **ARI/ILI/SARI** (n=45): 5 positive clinical/ spiked samples from each of the following diseases: 1. Influenza A virus @ 2. Influenza B virus @ 3. SARS-CoV-2 @ 4. RSV A/B @ 5. HPIV @ 6. HMPV @ 7. Adenovirus @ 8. Seasonal Coronaviruses * 9. Rhinovirus/Enterovirus* *Cross reactivity panel is arranged in descending order of priority.* *The pathogens marked @ are essentially to be tested.* *It is recommended to test for all pathogens listed in the cross reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*	Samples from cases of **AES** (n=35): 5 positive clinical/ spiked samples from each of the following diseases: 1. Japanese Encephalitis @ 2. Dengue @ 3. HSV @ 4. VZV @ 5. West Nile Virus * 6. Chandipura virus * 7. Rabies virus * *Cross reactivity panel is arranged in descending order of priority.* *The pathogens marked @ are essentially to be tested.* *It is recommended to test for all pathogens listed in the cross reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.*	5 positive clinical/ spiked samples from each of the following diseases, presenting with **respiratory and/or encephalitis symptoms** (n=20): 1. Measles 2. Rubella 3. Mumps 4. SARS-CoV-2
A.2.2 Samples from cases with acute respiratory disease/ acute encephalitis/ acute febrile illness and RT-PCR negative for the above-mentioned pathogens and Nipah virus	25	35	40

A.2.3 Healthy/ asymptomatic cases from endemic regions negative for Nipah virus	10	10	20
Serum/ throat swab/ urine samples collected from the same case may be used for evaluation. Commercially available validated standard panels that are accepted by accreditation agencies can also be used if the clinical sample with required target is not available for cross-reactivity analysis.			

10. Evaluation method:

The index test and the reference tests should be run simultaneously on the sample panel, and results should be recorded.

11. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

12. Resolution of discrepant results:

True positive samples: These are samples positive by reference assay and index test.
True negative samples: These are samples negative by reference assay and index test.
False positive samples: These are samples negative by reference assay and positive by index test.
False negative samples: These are samples positive by reference assay and negative by index test.

13. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

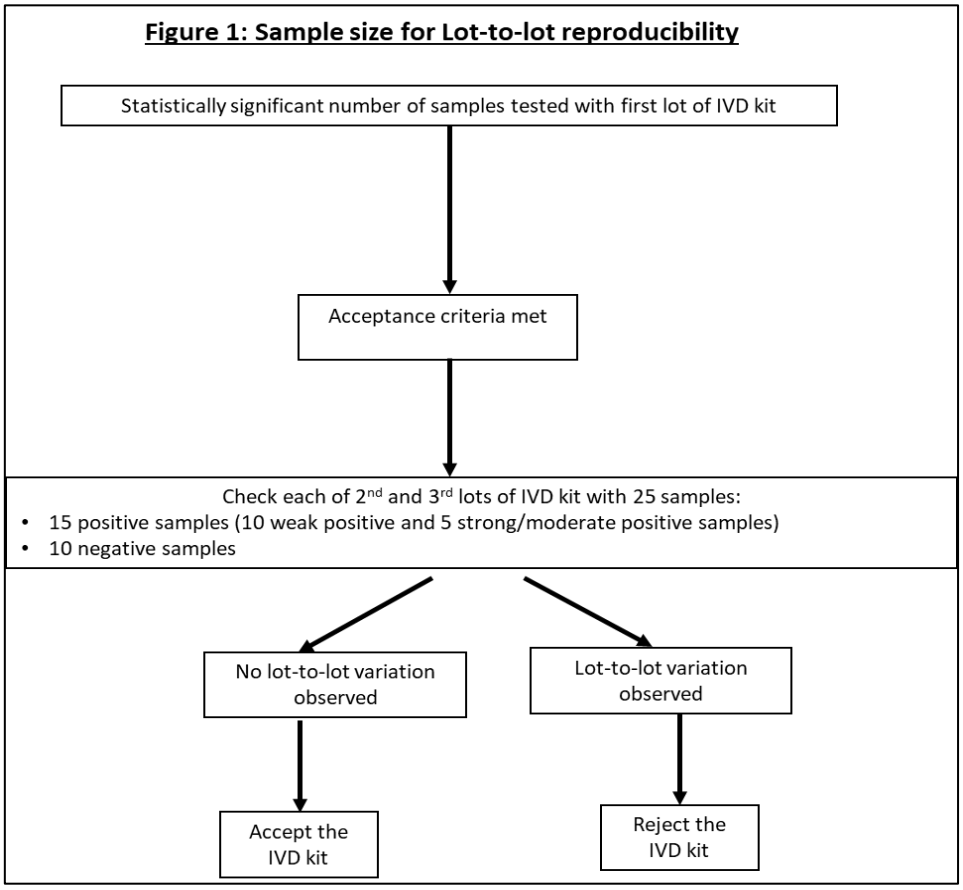
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Figure 1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs. Concordance should be 100% based on positive and negative test result (qualitative).

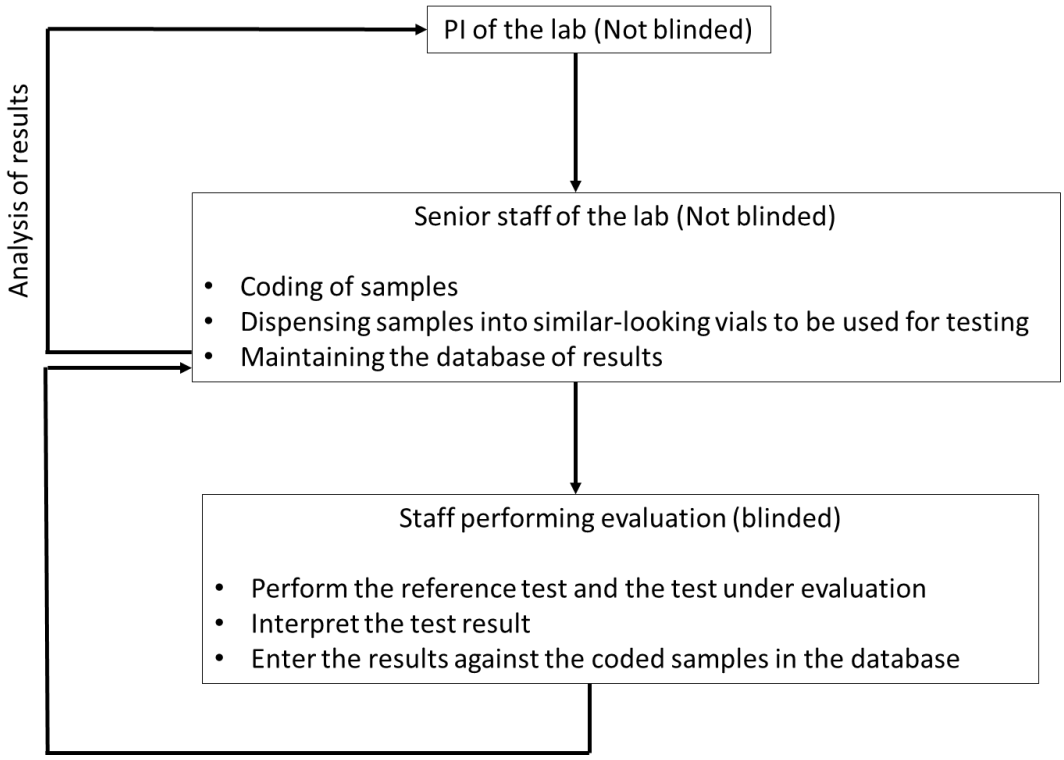
d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report Ct standard deviation across repeats (≤1cycle deviation recommended).

14. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



15. Acceptance Criteria

Sensitivity: ≥95%

Specificity: ≥98%

Cross reactivity with other viruses as outlined in the negative sample panel: Minimal

Invalid test rate: ≤5%

16. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References:

1. Yadav PD, Majumdar T, Gupta N, Kumar MA, Shete A, Pardeshi P, Sultana S, Sahay RR, Manoj MN, Patil S, Floura S, Gangakhedkar R, Mourya DT. Standardization & validation of Truenat™ point-of-care test for rapid diagnosis of Nipah. Indian J Med Res. 2021 Apr;154(4):645-649. doi: 10.4103/ijmr.IJMR_4717_20. PMID: 34854433; PMCID: PMC9205002.
2. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification – Diagnostic Assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1
3. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance evaluation report format

Performance evaluation report for Nipah virus real-time PCR kits

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: clinical/spiked, strong, moderate, weak/simulated samples)	
	Negative samples (provide details (clinical/spiked,), including cross reactivity panel/simulated samples)	

Results

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Nipah virus real-time PCR	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross reactivity
- Invalid test rate
- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ..... sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for Nipah virus…….. Kit (Lot No……) manufactured by …… (supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal …….

Annexure-1: Information on Operational and Test Performance Characteristics Required from Manufacturers

The manufacturer should provide the following details about the IVD:

1. Instructions for Use
2. Scope of the IVD:
3. Intended Use Statement
4. Principle of the assay
5. Intended testing population (cases of AES/ARI/SARI)
6. Intended user (laboratory professional and/or health care worker at point-of-care)
7. Lot/batch No.
8. Date of manufacture
9. Date of Expiry
10. Information on operational Characteristics
    i. Configuration of the kit/device
    ii. Requirement of any additional equipment, device
    iii. Requirement of any additional reagents
    iv. Operation conditions
    v. Storage and stability before and after opening
    vi. Internal control provided or not
    vii. Quality control and batch testing data
    viii. Biosafety aspects- waste disposal requirements
11. Information on Test Performance Characteristics
    i. Type of sample- Nasopharyngeal swab/Throat swab/ CSF/Serum / Other specimen
    ii. Volume of sample
    iii. Any specific sample NOT to be tested
    iv. Any additional sample processing required
    v. Any additional device/consumable like sample transfer device, pipette, tube, etc required
    vi. Name of analyte to be detected
    vii. Pathogen(s) targeted by the kit
    viii. Time taken for testing

ix. Time for result reading and interpretation

x. Manual or automated(equipment)reading

xi. Limit of detection/Limit of Quantification and range of detection

xii. Diagnostic sensitivity

xiii. Diagnostic specificity

xiv. Stability and reproducibility

xv. Training required for testing

xvi. If yes, duration

xvii. Details of Cut-off and /or Equivocal Zone for interpretation of test

xviii. Interpretation of invalid and indeterminate results to be provided

xix. It is recommended to provide data demonstrating accuracy and precision

*****Please mention "Not applicable" against sections not pertaining to the kit.

CHANDIPURA VIRUS REAL TIME PCR KIT

List of Contributors:

A. Working Group:

1. Dr. Vijay P Bondre, Scientist-G, ICMR-National Institute of Virology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
2. Ms. Krittika Bhattacharyya, Statistical Officer (Planning), Directorate of Economics and Statistics, Government of National Capital Territory of Delhi
3. Dr. Labanya Mukhopadhyay, Scientist-C, Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

B. Review Committee:

1. Dr. Vasanthapuram Ravi, Former Dean Research and Head of Neurovirology, National Institute of Mental Health and Neuro-Sciences, Bengaluru, Karnataka
2. Dr. Reeta Mani, Professor and Head of Neurovirology, National Institute of Mental Health and Neuro-Sciences, Bengaluru, Karnataka
3. Dr. Bhagirathi Dwibedi, Professor of Paediatrics, All India Institute of Medical Sciences, Bhubaneswar
4. Mr. Pramod Meshram, Deputy Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
5. Dr. Sella Senthil, Assistant Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
6. Dr. Md Omair Anwar, Drugs Inspector (Medical Devices), IVD Division, Central Drugs Standard Control Organization, Government of India
7. Dr. Nivedita Gupta, Scientist-G and Head of the Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

Performance evaluation protocol for Chandipura virus real-time PCR kits

I. Background

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding in-vitro diagnostic kit (IVD) performance.

This recommendation focuses on the laboratory performance evaluation of Chandipura virus (CHPV) virus real time PCR kit. All clinical samples tested in the study should be evaluated in accordance with the candidate test's instructions for use.

II. Purpose:

To evaluate the performance characteristics of CHPV real-time PCR kits in the diagnosis of CHPV infection/ disease using irreversibly de-identified leftover archived/ spiked clinical samples.

III. Requirements:

1. **Kits Under Evaluation**: Include detailed information such as batch number, lot number, expiry date, and other relevant specifications. For kits designed to operate within a closed system, manufacturers must provide the necessary equipment and consumables for testing.

2. **Evaluation Sites/Laboratories**: Identify laboratories equipped with the required instruments and infrastructure to conduct the evaluation.

3. **Reference Test Kits**: Use reference kits or in-house kits developed by the reference laboratory, which have been validated to demonstrate satisfactory performance.

4. **Evaluation Panel**: Prepare a panel of well-characterised clinical samples from confirmed cases or spiked samples for a comprehensive evaluation.

5. **Laboratory Supplies**: Ensure all necessary laboratory materials and supplies are available for the evaluation process.

IV. Ethical Approvals:

Performance evaluation activities using irreversibly de-identified clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. Study design/type: Diagnostic accuracy study using irreversibly de-identified archived clinical/spiked samples

2. Preparation of Evaluation sites/laboratories:

Identified IVD kit evaluation laboratories should establish their proficiency through the following:

A) Accreditation for at least one of the Quality management systems, such as

- Testing Laboratory or Calibration Laboratory (ISO/IEC 17025)
- Medical Laboratory (ISO 15189)
- Proficiency Testing Provider (ISO/IEC 17043)

  OR

- CDSCO-approved reference laboratory

B) Staff training: All staff involved in IVD kit evaluation process should undergo hands on training and competency assessment in the following areas:

- Preparation and characterization of kit evaluation panel
- Handling of Chandipura real-time PCR kits received for performance evaluation (verification/storage/unpacking etc.).
- Testing procedures, interpretation and recording of results, and reporting
- Data handling, data safety & confidentiality

3. Preparation of Chandipura RNA evaluation panel:

A well characterised panel of CHPV positive clinical samples is a critical requirement for evaluation of these RT-PCR IVD kits. A statistically significant number of clinical samples should be used for the evaluation.

The sample type for CHPV detection is Cerebrospinal fluid (CSF) and serum. If a kit claims to detect CHPV in both sample types, attempt should be made to evaluate the assay across both serum and CSF using statistically significant sample size for each sample type. In case all the sample types mentioned in the IFU are not available with the lab, the performance evaluation report should clearly mention the sample type against which the kit is evaluated, ensuring statistical rigor. There should be no ambiguity about the type of sample used for evaluation.

4. RNA extraction:

RNA extraction should be performed as per reference test and index test Instructions for Use (IFU).

If the manufacturer of the index test recommends a specific RNA extraction kit/system, the same needs to be provided by the manufacturer if the evaluating lab is unable to procure the same/it is not available within the lab's IVD evaluation scope.

5. Real-time PCR system:

PCR should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Real-time closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.

6. Internal Control/Extraction Control:

The index test must have an internal control (housekeeping gene), with or without an extraction control (RNA added before extraction to a sample).

7. Reference assay:

A WHO Pre-Qualified/ US FDA/ ATAGI Australia/ PMDA Japan approved real time CHPV PCR assay/ ICMR-National Institute of Virology, Pune developed protocol for detection of Chandipura virus RNA will serve as the reference assay.

All positive samples should be confirmed positive by the reference assay.

All negative samples should be confirmed negative by the reference assay and CHPV IgM.

8. Sample size for performance evaluation:

Sample size is calculated assuming 95% sensitivity and specificity of the index test, 95% confidence level, absolute precision of 5% and ≤5% invalid test rate. A minimum of 77 (rounded to 80) positive clinical samples and a minimum of 77 (rounded to 80) negative clinical samples for each sample type are required for performance evaluation. Sample sizes are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples.*
- *n (sp) is the minimum number of negative samples.*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96).*
- *Se is the predetermined sensitivity.*
- *Sp is the predetermined specificity.*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

9. Sample panel composition:

A) Positive samples (Minimum n=80 for each sample type): These samples should be clinical/spiked samples positive by reference real-time PCR assay and preferably represent all genetic variants. The distribution of samples should be as follows:

Characteristic of positive sample	Minimum no. of serum samples needed (for kits detecting CHPV in serum)	Minimum no. of CSF samples needed (for kits detecting CHPV in CSF)
A.1 Strong positive [Ct value ≤ 25]	24	24
A.2 Moderate positive [Ct value between >25 and ≤31]	28	28
A.3 Weak positive [Ct value >31 and ≤ 37]	28	28

For kits detecting CHPV in both serum and CSF, 80 positive serum samples and 80 positive CSF samples should be used for performance evaluation. One sample type should not be substituted by the other to reach the desired sample size in case there is paucity of samples.

Note: *Since such large number of positive clinical samples may **NOT** be available for Chandipura virus, pre-titrated and inactivated virus obtained from tissue culture fluid prepared in the laboratory will be used to spike serum and CSF samples [dilution factor: 1:10 to 1:1000 to generate samples with different intensities of positivity]. These spiked samples will be stored at -80°C, after being tested by the reference assay.*

B) Negative samples (n=80 for each sample type): All negative samples should be negative by reference assay and CHPV IgM. Distribution of the negative samples should be as follows:

Categories of samples as per the sample type	Sample type	
	Serum/plasma (Minimum n=80, (B.1 + B.2))	CSF (Minimum n=80, (B.1+B.2))
B.1 Samples from cases of AES/ spiked samples which are RT-PCR positive for known pathogens but negative for CHPV (CHPV RNA and serology)	**30** 5 positive clinical/ spiked samples from each of the following diseases (confirmed by PCR): 8. Dengue virus @ 9. Japanese Encephalitis @ 10. HSV 1/2 * 11. West Nile Virus* 12. VSV *	**35** 1. Seven (07) positive clinical/ spiked samples from each of the following diseases: a) Japanese Encephalitis @ b) Dengue virus @ c) HSV 1/2 * d) West Nile Virus * **17.** Rabies virus (n=4)* **18.** VSV (n=3)*

B.2 Samples from cases with acute encephalitis and RT-PCR negative for the above-mentioned pathogens and CHPV (CHPV RNA and serology)	**50**	**45**
B.3 Healthy/ asymptomatic cases from endemic regions negative for CHPV (CHPV RNA and serology)	5 (desirable, not mandatory)	20 (desirable, not mandatory)
Serum/plasma and CSF samples collected from the same case may be used for evaluation. *Cross reactivity panel is arranged in descending order of priority.* *The pathogens marked @ are essentially to be tested.* *It is recommended to test for all pathogens listed in the cross-reactivity panel. However, if there is an acute shortfall or non-availability of clinical samples, one may consider reducing only the pathogens of lower priority marked by * , while ensuring that the actual numbers of cross-reactive sample panel remain the same by compensating with the available "essentially to be tested" samples.* *Testing for Rabies and VSV is recommended since both the viruses belong to the same family as Chandipura virus (Rhabdoviridae). Spiked specimens/ synthetic transcripts may be used for these viruses. Commercially available validated standard panels that are accepted by accreditation agencies can also be used if clinical samplse with required target is not available.*		

10. Evaluation method:

The index test and reference tests should be conducted simultaneously on the sample panel to minimize the risk of false-negative results from the index test due to freeze-thaw cycles or sample degradation from prolonged storage.

11. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU.

12. Resolution of discrepant results:

True positive samples: These are samples positive by both the reference assay and index test.

True negative samples: These are samples negative by both the reference assay and index test.

False positive samples: These are samples negative by reference assay and positive by index test.

False negative samples: These are samples positive by reference assay and negative by index test.

13. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

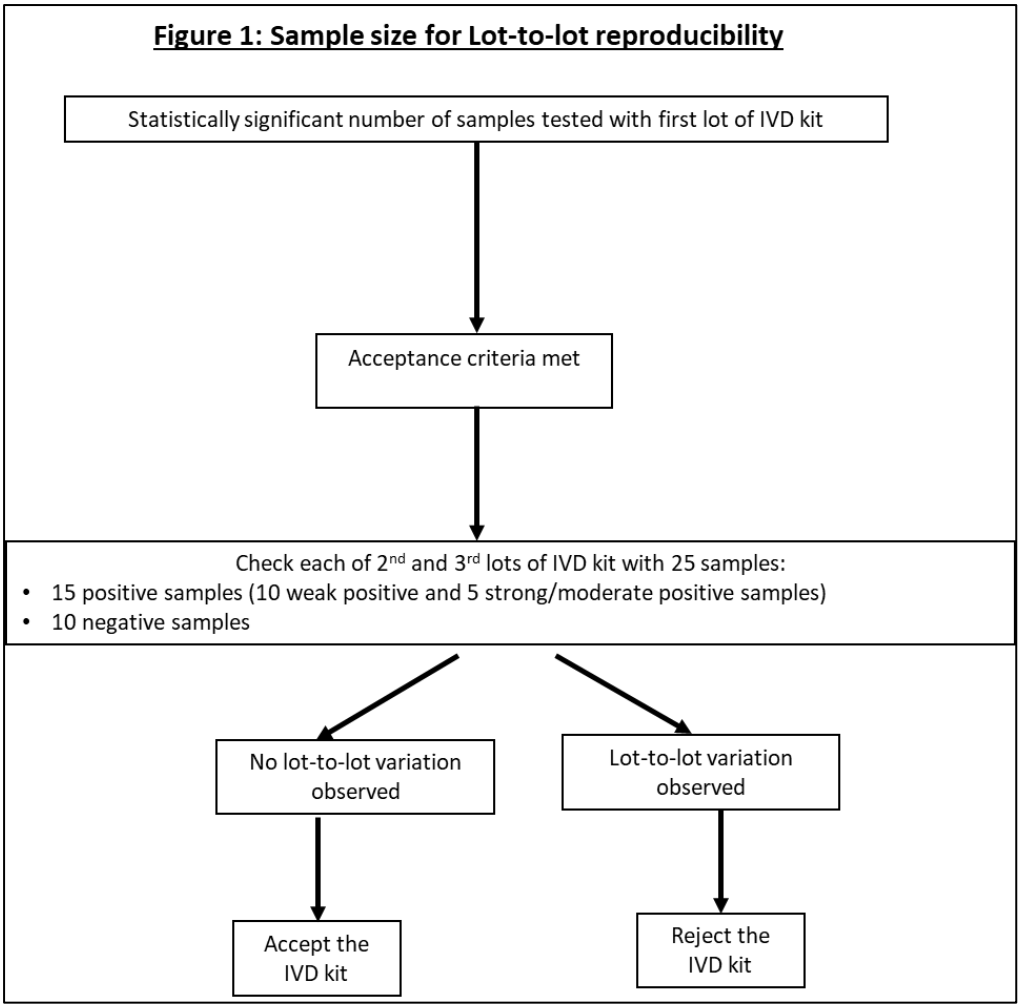
B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples comprising 10 low positive AND 5 moderate/strong positive samples, and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

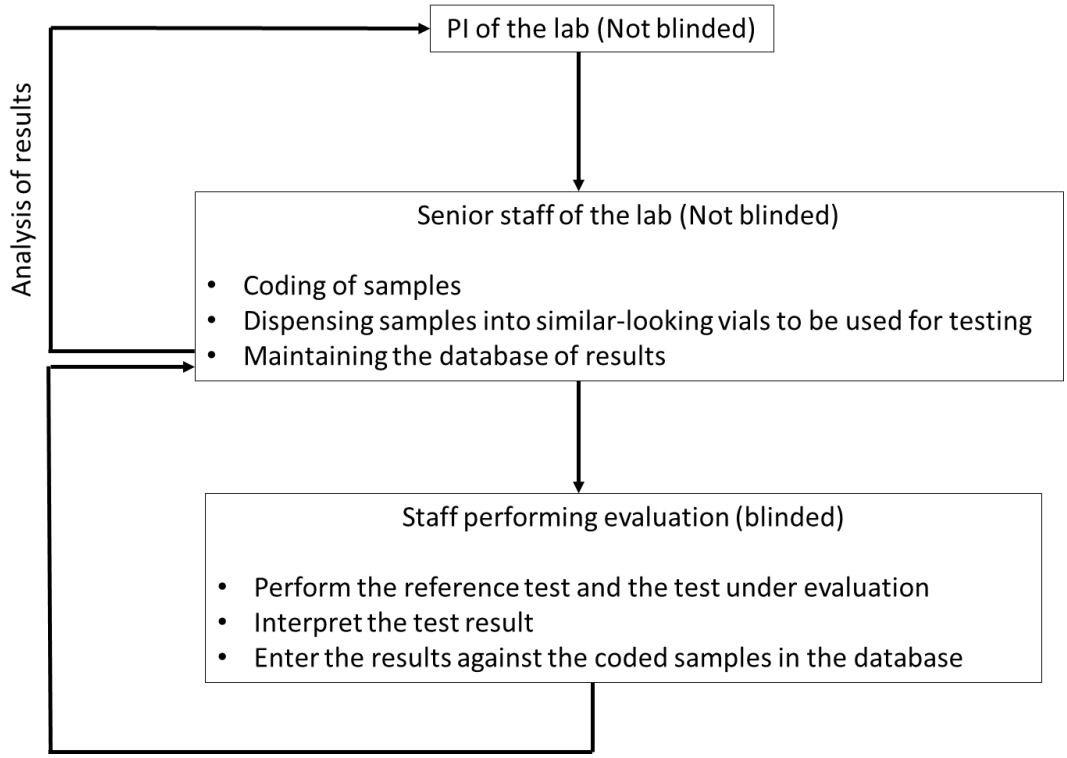
d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples (strong, moderate and weak positive samples) and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report Ct standard deviation across repeats (≤1cycle deviation recommended).

14. Blinding of laboratory staff

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should maintain confidentiality of data. Refer to Fig. 2.

Fig.2: Blinding in evaluation exercise



15. Acceptance criteria:

Sensitivity: ≥ 95%

Specificity: ≥ 98%

Cross-reactivity with other rhabdoviruses: Minimal

Invalid test rate ≤5%

16. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the field evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

VI. References

1. Sudeep AB, Gurav YK, Bondre VP. Changing clinical scenario in Chandipura virus infection. *Indian J Med Res*. 2016;143(6):712-721. doi:10.4103/0971-5916.191929.

2. Sapkal GN, Sawant PM, Mourya DT. Chandipura Viral Encephalitis: A Brief Review. Open Virol J. 2018 Aug 31;12:44-51. doi: 10.2174/1874357901812010044. PMID: 30288194; PMCID: PMC6142667.

3. World Health Organization. Technical Guidance Series (TGS) for WHO Prequalification-Diagnostic assessment TGS-3. 2017. Available at: https://iris.who.int/bitstream/handle/10665/258985/WHO-EMP-RHT-PQT-TGS3-2017.03-eng.pdf;sequence=1

4. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

VII. Performance Evaluation Report Format

Performance evaluation report for Chandipura virus real-time PCR kits

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number:Issue date: Valid Up to:		
Application No.		
Sample Panel	Positive samples (provide details: type,strong, moderate, weak/simulated samples)	
	Negative samples (provide details, type,including cross reactivity panel/simulated samples)	

Results

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of Chandipura real-time PCR kits	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross reactivity with related viruses:
- Invalid test rate:
- **Performance: Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ….. sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for Chandipura…….. Kit (Lot No……) manufactured by …… (supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……

Annexure-1: Information on Operational and Test Performance Characteristics Required from Manufacturers

The manufacturer should provide the following details about the IVD:

1. Instructions for Use
2. Scope of the IVD:
3. Intended Use Statement
4. Principle of the assay
5. Intended testing population (cases of Acute Febrile Illness/ AES)
6. Intended user (laboratory professional and/or health care worker at point-of-care)
7. Lot/batch No.
8. Date of manufacture
9. Date of Expiry
10. Information on operational Characteristics
    i. Configuration of the kit/device
    ii. Requirement of any additional equipment, device
    iii. Requirement of any additional reagents
    iv. Operation conditions
    v. Storage and stability before and after opening
    vi. Internal control provided or not
    vii. Quality control and batch testing data
    viii. Biosafety aspects- waste disposal requirements
11. Information on Test Performance Characteristics
    i. Type of sample-CSF/Serum/Other specimen
    ii. Volume of sample
    iii. Any specific sample NOT to be tested
    iv. Any additional sample processing required
    v. Any additional device/consumable like sample transfer device, pipette, tube, etc required
    vi. Name of analyte to be detected
    vii. Pathogen(s) targeted by the kit

viii. Time taken for testing

ix. Time for result reading and interpretation

x. Manual or automated (equipment) reading

xi. Limit of detection/Limit of Quantification and range of detection

xii. Diagnostic sensitivity

xiii. Diagnostic specificity

xiv. Stability and reproducibility

xv. Training required for testing

xvi. If yes, duration

xvii. Details of Cut-off and /or Equivocal Zone for interpretation of test

xviii. Interpretation of invalid and indeterminate results to be provided

xix. It is recommended to provide data demonstrating accuracy and precision

xx. Limit of detection

*****Please mention "Not applicable" against sections not pertaining to the kit.

TYPHOID IN-VITRO DIAGNOSTICS

List of Contributors:

A. Working Group:

1. Dr. Rajlakshmi Viswanathan, Scientist-E, ICMR-National Institute of Virology, Department of Health Research, Ministry of Health and Family Welfare, Government of India
2. Dr. Labanya Mukhopadhyay, Scientist-C, Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India
3. Ms. Krittika Bhattacharyya, Statistical Officer (Planning), Directorate of Economics and Statistics, Government of National Capital Territory of Delhi
4. Dr. Suman Kanungo, Scientist-G and Director, ICMR-Regional Medical Research Center, Dibrugarh, Department of Health Research, Ministry of Health and Family Welfare, Government of India
5. Dr. Archana Angrup, Additional Professor, Department of Medical Microbiology, Postgraduate Institute of Medical Education and Research, Chandigarh
6. Dr Shafeeq K Shahul Hameed, Former Scientific Officer, Center for Brain Research, Indian Institute of Science, Bengaluru, and currently Senior Epidemiologist, Health Emergency Department, Ministry of Public Health, Qatar
7. Dr. Pratibha Kale, Additional Professor, Department of Clinical Microbiology, Institute of Liver and Biliary Sciences, New Delhi
8. Dr Jeromie Wesley Vivian T, Scientist C, ICMR-National Institute of Epidemiology, Department of Health Research, Ministry of Health and Family Welfare, Government of India

B. Review Committee:

1. Dr. Vasanthapuram Ravi, Former Dean Research and Head of Neurovirology, National Institute of Mental Health and Neuro-Sciences, Bengaluru, Karnataka
2. Dr. Arti Kapil, Professor & Head Department of Microbiology, North DMC Medical College and Hindu Rao Hospital, New Delhi & Ex-Professor and In-Charge of Bacteriology at All India Institute of Medical Sciences, Delhi
3. Dr. V Balaji, Professor of Microbiology, Christian Medical College, Vellore India
4. Dr. Jacob John, Professor of Community Health, Christian Medical College, Vellore India
5. Mr. Pramod Meshram, Deputy Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
6. Dr. Sella Senthil, Assistant Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
7. Dr. Md Omair Anwar, Drugs Inspector (Medical Devices), IVD Division, Central Drugs Standard Control Organization, Government of India
8. Dr. Nivedita Gupta, Scientist-G and Head of the Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

Field Evaluation Protocol for Typhoid Molecular IVDs

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Typhoid molecular IVD kits in the diagnosis of Typhoid fever in individuals with unknown disease status.

The protocol outlines field evaluation of IVD kits that detect S Typhi gene segments in blood/ serum/plasma samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Laboratory supplies

IV. Ethical approval:

The study will be initiated after obtaining approval from the institutional human ethics committee.

V. Procedure:

1. **Study design/type**: Cross-sectional study
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through

   A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider (ISO/IEC 17043) or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel
   - Handling of Typhoid molecular IVD received for performance evaluation (Verification/Storage/Unpacking etc).
   - Testing, interpreting, recording of results & reporting
   - Data handling, data safety & confidentiality

3. Sample size for performance evaluation:

Sample size has been calculated assuming 95% level of significance, an absolute precision of 5%, sensitivity of ≥95%, specificity of ≥95%, invalid test rate ≤5%. Sample size is calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR}) xP}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR}) xP}$$

- n (se) is the minimum number of positive samples
- n (sp) is the minimum number of negative samples
- Z2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z2 =1.96)
- Se is the predetermined sensitivity
- Sp is the predetermined specificity
- d is the predetermined marginal error (5%)
- IR is the invalid test rate
- P is the prevalence of the disease

A minimum of 2000 cases with acute febrile illness satisfying the case definition need to be enrolled in the study to achieve statistically significant minimum number of positive samples.

Explanation: A minimum of 77 positive cases and a minimum of 77 negative cases should be enrolled as per the aforementioned formulae. Rounding up, a minimum of 80 positive cases and a minimum of 80 negative cases should be enrolled for evaluation. Since approximately 4% of acute fever cases are positive for typhoid through blood culture (John J, 2023), approximately 2000 cases satisfying the case definition need to be enrolled to achieve the desired target of 80 positive cases. This sample size is adequate for the desired number of negative cases that need to be tested in the evaluation.

Recruitment of cases shall be halted once desired number of positive and negative samples are reached.

Note: Typhoid being a seasonal disease, attempts should be made to enrol more cases during seasons of high transmission. For multi-centric studies, clinicians at each participating site should be sensitized on case enrolment.

4. Inclusion criteria:

The case definition from India's Integrated Disease Surveillance Programme (IDSP) is adapted to increase the probability of enrolment of blood culture positive typhoid fever,

and ensure availability of adequate sample volumes (Ministry of Health & Family Welfare, 2024). Briefly, an individual aged 2-65 years from OPD/IPD settings will be included if they satisfy both the major and any one of the minor criteria mentioned below:

Major criteria: Documented fever of >38°C AND Duration of fever 4-10 days (preferably ≥3 consecutive days of fever)

Minor criteria:

- Toxic look
- Coated tongue
- Relative bradycardia
- Splenomegaly

5. Exclusion criteria:

- Age <2 years or ≥65 years
- Pregnancy
- Chronic diseases
- Known or identified focus of infection

6. Nucleic acid extraction

Nucleic acid extraction should be performed as per reference test and index test (kit under evaluation) Instructions for Use (IFU).

If the manufacturer of the index test recommends a specific nucleic acid extraction kit/system, the same needs to be provided by the manufacturer if the evaluating lab is unable to procure the same/it is not available within the lab's IVD evaluation scope.

7. IVD System (e.g.: Real-Time PCR System)

The reference test and index test should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.

8. Internal control/Extraction control

The test under evaluation (index test) must have an internal control (housekeeping gene), with or without an extraction control (nucleic acid added before extraction to a sample).

9. Reference standard:

- A composite gold standard (blood culture **AND/OR** accredited molecular test/WHO Pre-Qualified/US FDA/ ATAGI Australia/ PMDA Japan approved molecular test – e.g.: for *stag* and/or *vi* genes) should be used as the reference standard. Use of composite gold standard is intended to increase the ability to detect a true-positive case.
  - *Blood culture alone should be used as the reference standard till the time a well-performing reference molecular test (as outlined in this section) is available with the evaluating labs.*
  - *Automated blood culture and ID methods should be used, and these tests should be under laboratory accreditation scope.*
  - *Note: WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Typhoid molecular kit may be used as reference molecular test as and when these kits become available.*

10. Study implementation:

The individuals satisfying the case definition will be recruited, and sample(s) will be collected as per the requirements of the study. The samples should be tested by the reference standard and the index test.

11. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU/SoP.

12. Cross-reactivity: Cross-reactivity should be reported separately using the samples mentioned below:

Cross-reactivity panel[@]:	**Total 30 samples (described below)**
-Acute phase samples from proven Paratyphoid A, B, NTS infections -Acute phase samples from confirmed cases of Brucellosis, Rickettsial infections, Enterobacteriaceae infections (e.g. sepsis), and Influenza	**1. <u>Paratyphoid samples (n=10)</u>** • Paratyphoid A (n=4) • Paratyphoid B (n=3) • Paratyphoid C (n=3) 2. **<u>NTS samples (n=10)</u>** One each for the following 10 NTS[##]: • Salmonella Typhimurium • Salmonella Enteritidis • Salmonella Kentucky • Salmonella Eastbourne • Salmonella Dublin • Salmonella Bareilly • Salmonella Weltevreden

	• Salmonella Newport • Salmonella Infantis • Salmonella Agona **3.<u>Cases of Brucellosis and Rickettsial infections (n=5)</u>** • Brucellosis (B. abortus and/or *B. melitensis* and/or *B. suis*) cases (n=2) * • Rickettsial (*R. rickettsiae* and/or *R. conorii* and/or *R. typhi*) infections (n=3) * **4. <u>Enterobacteriaceae infection (n=3)</u>** Cases of Enterobacteriaceae infection (sepsis) – including E. coli, *Klebsiella sp.*, *Citrobacter sp.* **5. <u>Influenza cases (n=2)</u>** 2 cases of seasonal influenza virus infection (including influenza A and B)

NTS = Non Typhoidal Salmonella

@ These samples should be characterized as per national guidelines/globally acceptable standards

In case of unavailability of samples positive for a cross-reacting pathogen, the gap can be met with negative samples (negative for typhoid markers and other causes of acute febrile illness such as Dengue, malaria, chikungunya, leptospira, scrub typhus) spiked with requisite serovars. It is recommended to use well-characterized serovars from reputed and credible national/global sources or agencies.

* Commercially available validated panels containing genomes of these pathogens, that are accepted by accreditation agencies, may be used in case of paucity of clinical samples.

13. Statistical analysis:

Sensitivity and specificity will be calculated.

Interim analysis of data shall be conducted on completing evaluation of 25%, 50% and 75% of samples. If, at any point, the performance of the assay is found to be not satisfactory, the assay shall not be evaluated further. Evaluation fee shall be charged accordingly.

14. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples and 3 negative samples should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

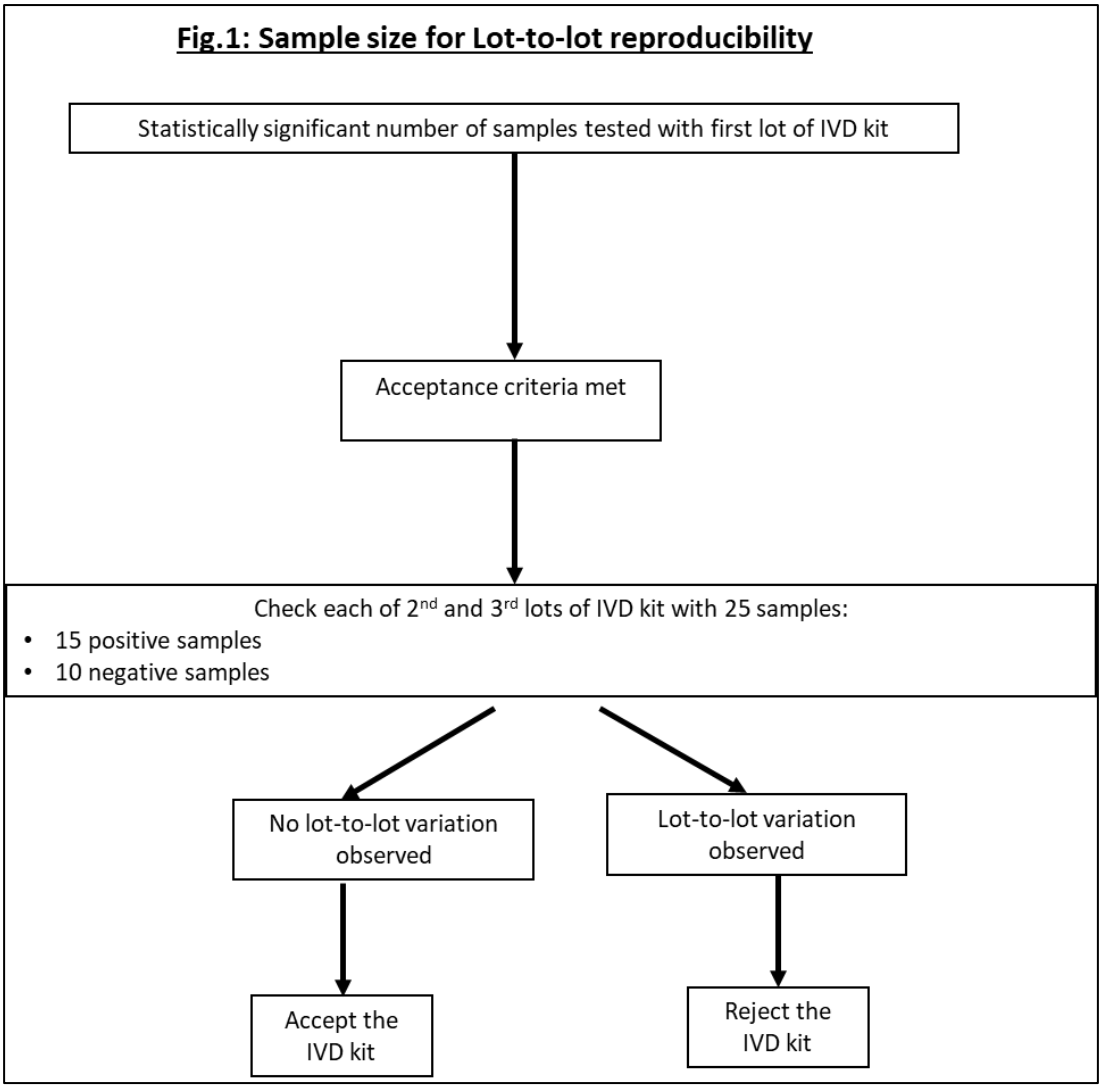
B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report standard deviation of measured value across repeats (recommended ≤1 standard deviation variation).

15. Resolution of discrepant results:

True positive samples: These are samples positive by reference standard and index test.
True negative samples: These are samples negative by reference standard and index test.
False positive samples: These are samples negative by reference standard and positive by index test.
False negative samples: These are samples positive by reference standard and negative by index test.

16. Blinding of laboratory staff

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the results of the reference assay. The PI of the evaluation exercise should remain unblinded, i.e., privy to the results of the reference test. Another senior laboratory staff selected by the PI may remain unblinded for overseeing the activity and maintaining the database of results. The data should be analyzed only by the PI of the evaluating lab. The PI should ensure data confidentiality.

17. Acceptance Criteria

Sensitivity: ≥95%
Specificity: ≥95%
Cross-reactivity with other pathogens listed in the negative sample panel: Minimal
Invalid test rate: ≤5%

18. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.
Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.
Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

References:

1. John J, Bavdekar A, Rongsen-Chandola T, Dutta S, Gupta M, Kanungo S, et al. Burden of typhoid and paratyphoid fever in India. N Engl J Med 2023;388(16): 1491–500. https://doi.org/10.1056/NEJMoa2209449. PMID: 37075141; PMCID: PMC10116367.
2. Ministry of Health and Family Welfare-Integrated Disease Surveillance Programme. Training manual for state & district surveillance officers case definitions of diseases & syndromes under surveillance module -5. https://idsp.mohfw.gov.in/WriteReadData/OldSite/2WkDSOSept 08/Resources_files/DistrictSurvMan/Module5.pdf. [Accessed 23 November 2024].
3. Viswanathan, R., Mukhopadhyay, L., Angrup, A., Kale, P., Bhattacharyya, K., Kanungo, S., Jeromie Wesley, V. T., Shahul Hameed, S. K., Sella Senthil, M., Meshram, P., Kapil, A., Balaji, V., John, J., Ravi, V., & Gupta, N. (2025). Standard protocols for performance evaluation of typhoid fever in-vitro diagnostic assays in laboratory and field settings. *Indian journal of medical microbiology*, *55*, 100857. https://doi.org/10.1016/j.ijmmb.2025.100857
4. Mather RG, Hopkins H, Parry CM, Dittrich S. Redefining typhoid diagnosis: what would an improved test need to look like? BMJ Glob Health 2019;4(5):e001831. https://doi.org/10.1136/bmjgh-2019-001831. PMID: 31749999; PMCID: PMC6830052.
5. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

Performance Evaluation Protocol for Typhoid Molecular IVDs

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured diagnostic kits appropriate for use in India. Hence the following guidelines shall establish uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Typhoid molecular IVD kits in the diagnosis of Typhoid fever using irreversibly de-identified leftover archived clinical samples.

The protocol outlines performance evaluation of IVD kits that detect S Typhi gene segments in blood/ serum/plasma samples.

Note: Performance evaluation with well-characterized leftover samples will be feasible once a well-performing reference molecular test (as outlined later in the document) and requisite clinical samples are available for use.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment, and blood culture facilities)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.
Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through

A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider (ISO/IEC 17043) or CDSCO approved Reference laboratory.

B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

- Preparation & characterization of kit evaluation panel
- Handling of kits received for performance evaluation (Verification/Storage/Unpacking etc).
- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Typhoid molecular IVD evaluation panel

Well characterised sample panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of samples across relevant sample matrices should be available from cases with confirmed typhoid fever status.

4. Nucleic acid extraction

Nucleic acid extraction should be performed as per reference test and index test (kit under evaluation) Instructions for Use (IFU).

If the manufacturer of the index test recommends a specific nucleic acid extraction kit/system, the same needs to be provided by the manufacturer if the evaluating lab is unable to procure the same/it is not available within the lab's IVD evaluation scope.

5. IVD System (e.g.: Real-Time PCR System)

The reference test and index test should be performed using IVD-approved machines. If any equipment(s) is specified in the IFU of the index test, it should be used for the evaluation, and it should be provided by the manufacturer if not available within the lab's IVD evaluation scope.

Closed systems/devices awaiting evaluation should be provided by the manufacturer along with all necessary components, supplies and reagents.

6. Internal control/Extraction control

The test under evaluation (index test) must have an internal control (housekeeping gene), with or without an extraction control (nucleic acid added before extraction to a sample).

7. Reference standard:

- A composite gold standard (blood culture **AND/OR** accredited well-performing molecular test/WHO Pre-Qualified/US FDA/ ATAGI Australia/ PMDA Japan approved molecular test – e.g.: for *stag* and/or *vi* genes) should be used to characterize the samples. Use of composite gold standard is intended to increase the ability to detect a true-positive sample.
    - *Automated blood culture and ID methods should be used, and these tests should be under laboratory accreditation scope.*

It should be noted that the samples for blood culture and molecular tests may be different. It is recommended to build a biorepository of well-characterized positive and negative samples across relevant sample matrices from confirmed positive and negative cases of typhoid fever.

8. Sample size and sample panel composition: Sample size has been calculated assuming 95% level of significance, an absolute precision of 5%, sensitivity of ≥95%, specificity of ≥95%, and invalid test rate ≤5%. Sample sizes are calculated using the following formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples*
- *n (sp) is the minimum number of negative samples*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96)*
- *Se is the predetermined sensitivity*
- *Sp is the predetermined specificity*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

A minimum of 77 positive samples and a minimum of 77 negative samples should be used. Rounding up, a minimum of 80 positive samples and a minimum of 80 negative samples should be used for evaluation.

A. Positive samples: These samples should be acute phase samples from confirmed cases of typhoid fever positive by composite gold standard, but negative for common pathogens causing febrile illness pathogens (Dengue, Malaria, Scrub Typhus, Leptospira, Chikungunya).
B. Negative samples: These samples should be negative by both blood culture **AND** reference molecular test (once the requisite molecular test is available for use in India)

The sample size and sample panel composition are depicted in Table 1:

Table 1: Sample size and panel composition for performance evaluation of molecular IVD kits:

A. Minimum Positive samples in the panel (N= 80)	**Minimum Number of samples needed**
Composite gold standard positive acute phase samples which are negative for common febrile illness pathogens in India[@] (dengue, malaria, scrub typhus, leptospira, chikungunya)	**80**
B. Minimum Negative samples in the panel (N=80) *(samples negative by both blood culture **AND** reference molecular test)*	**Minimum Number of samples needed**

Samples from afebrile cases where blood culture and reference molecular test are negative for typhoid (e.g., healthy blood donors) *Afebrile: absence of fever for last 14 days*	**30**
Acute phase samples from non-typhoid febrile cases– positive for dengue, malaria, scrub typhus, leptospira, chikungunya @	**20** (equal distribution for these diseases)
Cross-reactivity panel[@]: -Acute phase samples from proven Paratyphoid A, B, NTS infections -Acute phase samples from confirmed cases of Brucellosis, Rickettsial infections, Enterobacteriaceae infections (e.g. sepsis), and Influenza	**Total 30 samples (described below)** **2. <u>Paratyphoid samples (n=10)</u>** • Paratyphoid A (n=4) • Paratyphoid B (n=3) • Paratyphoid C (n=3) 2. **<u>NTS samples (n=10)</u>** One each for the following 10 NTS[##]: • Salmonella Typhimurium • Salmonella Enteritidis • Salmonella Kentucky • Salmonella Eastbourne • Salmonella Dublin • Salmonella Bareilly • Salmonella Weltevreden • Salmonella Newport • Salmonella Infantis • Salmonella Agona **3.<u>Cases of Brucellosis and Rickettsial infections (n=5)</u>** • Brucellosis (B. abortus and/or *B. melitensis* and/or *B. suis*) cases (n=2) * • Rickettsial (*R. rickettsiae* and/or *R. conorii* and/or *R. typhi*) infections (n=3) * **4. <u>Enterobacteriaceae infection (n=3)</u>** Cases of Enterobacteriaceae infection (sepsis) – including E. coli, *Klebsiella sp.*, *Citrobacter sp.* **6. <u>Influenza cases (n=2)</u>** cases of seasonal influenza virus infection (including influenza A and B)

NTS = Non Typhoidal Salmonella
@ These samples should be characterized as per national guidelines/globally acceptable standards

In case of unavailability of samples positive for a cross-reacting pathogen, the gap can be met with negative samples (negative for typhoid markers and other causes of acute febrile illness such as Dengue, malaria, chikungunya, leptospira, scrub typhus) spiked with requisite serovars. It is recommended to use well-characterized serovars from reputed and credible national/global sources or agencies.

* Commercially available validated panels containing genomes of these pathogens, that are accepted by accreditation agencies, may be used in case of paucity of clinical samples.

8. Evaluation method:

It is recommended to run the reference molecular test and the index test in parallel.

9. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU/SoP.

10. Resolution of discrepant results:

True positive samples: These are samples positive by reference test and index test.
True negative samples: These are samples negative by reference test and index test.
False positive samples: These are samples negative by reference test and positive by index test.
False negative samples: These are samples positive by reference test and negative by index test.

12. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

Reproducibility testing should include the following:

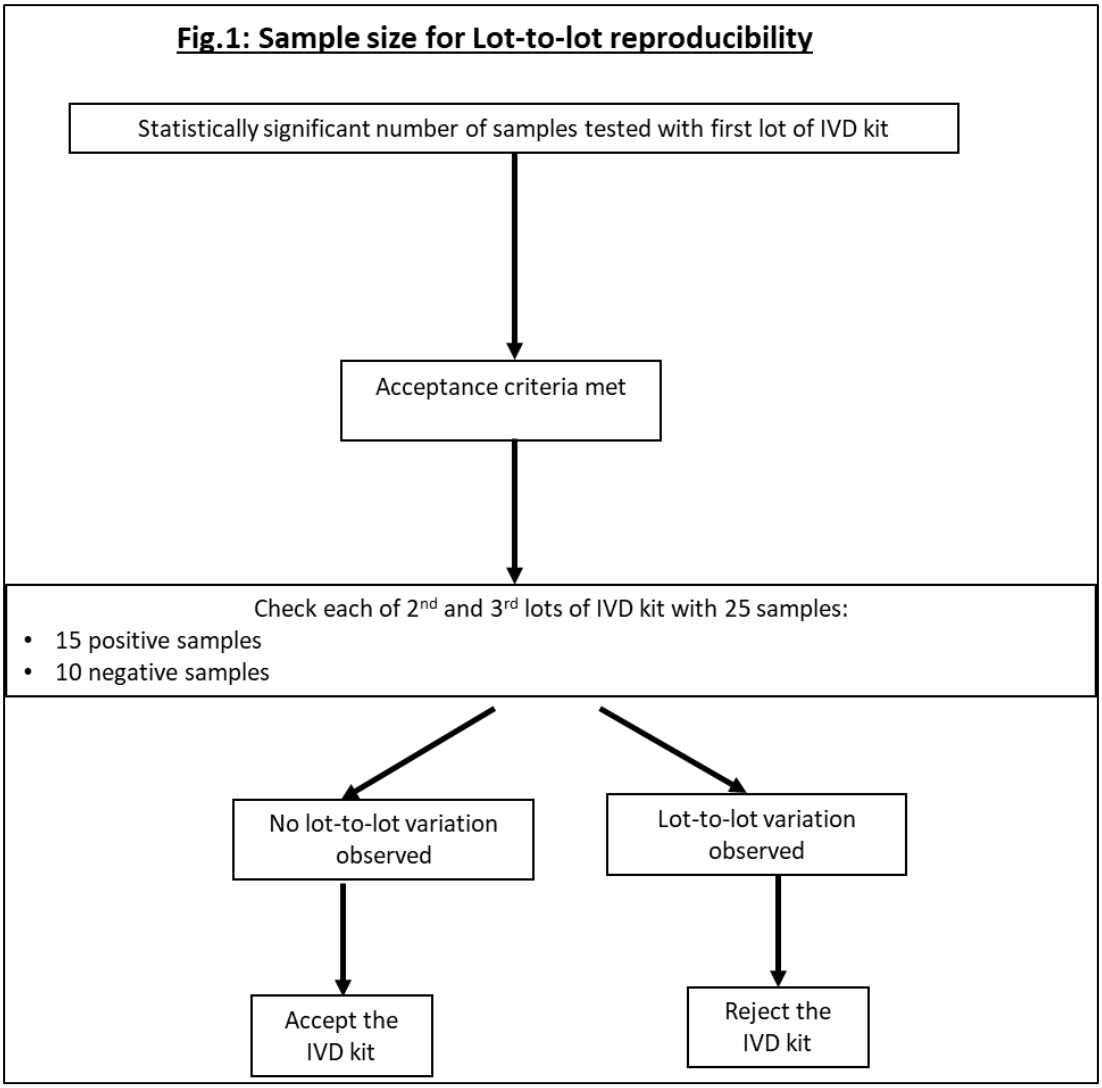
a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.

- Second lot of the kit: should be tested on 25 samples (15 positive samples and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Fig.1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

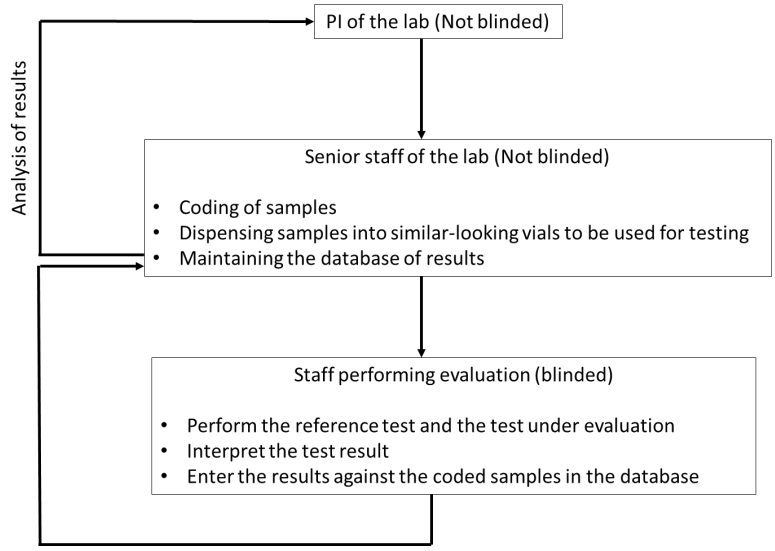
c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

It is desirable to report standard deviation of measured value across repeats (recommended ≤1 standard deviation variation).

13. Blinding of laboratory staff

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should ensure confidentiality of the data. Refer to Fig. 1.
Fig.1: Blinding in evaluation exercise



14. Acceptance Criteria

Sensitivity: ≥95%
Specificity: ≥95%
Cross-reactivity with other pathogens listed in the negative sample panel: Minimal
Invalid test rate: ≤5%

15. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

References:

1. John J, Bavdekar A, Rongsen-Chandola T, Dutta S, Gupta M, Kanungo S, et al. Burden of typhoid and paratyphoid fever in India. N Engl J Med 2023;388(16): 1491–500. https://doi.org/10.1056/NEJMoa2209449. PMID: 37075141; PMCID: PMC10116367.
2. Ministry of Health and Family Welfare-Integrated Disease Surveillance Programme. Training manual for state & district surveillance officers case definitions of diseases & syndromes under surveillance module -5. https://idsp.mohfw.gov.in/WriteReadData/OldSite/2WkDSOSept 08/Resources_files/DistrictSurvMan/Module5.pdf. [Accessed 23 November 2024].
3. Viswanathan, R., Mukhopadhyay, L., Angrup, A., Kale, P., Bhattacharyya, K., Kanungo, S., Jeromie Wesley, V. T., Shahul Hameed, S. K., Sella Senthil, M., Meshram, P., Kapil, A., Balaji, V., John, J., Ravi, V., & Gupta, N. (2025). Standard protocols for performance evaluation of typhoid fever in-vitro diagnostic assays in laboratory and field settings. *Indian journal of medical microbiology*, *55*, 100857. https://doi.org/10.1016/j.ijmmb.2025.100857
4. Mather RG, Hopkins H, Parry CM, Dittrich S. Redefining typhoid diagnosis: what would an improved test need to look like? BMJ Glob Health 2019;4(5):e001831. https://doi.org/10.1136/bmjgh-2019-001831. PMID: 31749999; PMCID: PMC6830052.
5. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

Field Evaluation Protocol for Typhoid Antigen-based IVDs

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostics kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Typhoid antigen-based kits in the diagnosis of Typhoid fever in individuals with unknown disease status.

The protocol outlines field evaluation of IVD kits that detect S Typhi antigens in blood/ serum/plasma samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Laboratory supplies

IV. Ethical approval:

The study will be initiated after obtaining approval from the institutional human ethics committee.

V. Procedure:

1.Study design/type: Cross-sectional study

2.Preparation of Evaluation sites/laboratories:

Identified IVD kit evaluation laboratories should establish their proficiency through

A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider (ISO/IEC 17043) or CDSCO approved Reference laboratory.

B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

- Preparation & characterization of kit evaluation panel
- Handling of Typhoid antigen-based IVD received for performance evaluation

(Verification/Storage/Unpacking etc).

- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Sample size for performance evaluation:

Sample size has been calculated assuming 95% level of significance, an absolute precision of 5%, sensitivity of ≥90%, specificity of ≥95%, and invalid test rate ≤5% (where applicable). Sample size is calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR}) xP}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ \left(1 - S_p\right)}{d^2 \text{ x } (1 - \text{IR}) xP}$$

- *n (se) is the minimum number of positive samples*
- *n (sp) is the minimum number of negative samples*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96)*
- *Se is the predetermined sensitivity*
- *Sp is the predetermined specificity*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*
- *P is the prevalence of the disease*

A minimum of 3750 cases with acute febrile illness satisfying the case definition need to be enrolled in the study to achieve statistically significant minimum number of positive samples.

Explanation: A minimum of *146 positive cases (139 samples for assays without the provision of invalid test)* and a minimum of *77 negative cases (73 samples for assays without the provision of invalid test)* should be enrolled as per the aforementioned formulae. Rounding up, a minimum of 150 positive cases and a minimum of 80 negative cases should be enrolled for evaluation. Since approximately 4% of acute fever cases are positive for

typhoid through blood culture (John J, 2023), approximately 3750 cases satisfying the case definition need to be enrolled to achieve the desired target of 150 positive cases. This sample size is adequate for the desired number of negative cases that need to be tested in the evaluation.

Recruitment of cases shall be halted once desired number of positive and negative samples are reached.

Note: Typhoid being a seasonal disease, attempts should be made to enrol more cases during seasons of high transmission. For multi-centric studies, clinicians at each participating site should be sensitized on case enrolment.

4. Inclusion criteria:

The case definition from India's Integrated Disease Surveillance Programme (IDSP) was adapted to increase the probability of enrolment of typhoid blood culture positive, and ensure availability of adequate sample volumes (Ministry of Health & Family Welfare, 2024). Briefly, an individual aged 2-65 years from OPD/IPD settings will be included if they satisfy both the major and any one of the minor criteria mentioned below:

Major criteria: Documented fever of >38°C **AND** Duration of fever 4-10 days (preferably ≥3 consecutive days of fever)

Minor criteria:

- Toxic look
- Coated tongue
- Relative bradycardia
- Splenomegaly

5. Exclusion criteria:

- Age <2 years or ≥65 years
- Pregnancy
- Chronic diseases
- Known or identified focus of infection

6. Reference standard:

- A composite gold standard (blood culture **AND/OR** accredited molecular test/WHO Pre-Qualified/US FDA/ ATAGI Australia/ PMDA Japan approved molecular test – e.g.: for *stag* and/or *vi* genes) should be used as the reference standard. Use of composite gold standard is intended to increase the ability to detect a true-positive case.
    - *Blood culture alone should be used as the reference standard till the time a well-performing reference molecular test (as outlined in this section) is available with the evaluating labs.*
    - *Automated blood culture and ID methods should be used, and these tests should be under laboratory accreditation scope.*

- *Note: WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Typhoid molecular kit may be used as reference molecular test as and when these kits become available.*

7. Study implementation:

The individuals satisfying the case definition will be recruited, and sample(s) will be collected as per the requirements of the study. The requisite samples should be tested by the reference standard and the index test.

8. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU/SoP.

9. Cross-reactivity: Cross-reactivity should be reported separately using the samples mentioned below:

Cross-reactivity panel[@]: -Acute phase samples from proven Paratyphoid A, B, NTS -Acute phase samples from confirmed cases of Brucellosis, Rickettsial infections, Enterobacteriaceae infections (e.g. sepsis), and Influenza	**Total 30 samples (described below)** **1. <u>Paratyphoid samples (n=10)</u>** • Paratyphoid A (n=4) • Paratyphoid B (n=3) • Paratyphoid C (n=3) 2. **<u>NTS samples (n=10)</u>** One each for the following 10 NTS[##]: • Salmonella Typhimurium • Salmonella Enteritidis • Salmonella Kentucky • Salmonella Eastbourne • Salmonella Dublin • Salmonella Bareilly • Salmonella Weltevreden • Salmonella Newport • Salmonella Infantis • Salmonella Agona **3.<u>Cases of Brucellosis and Rickettsial infections (n=5)</u>** • Brucellosis (B. abortus and/or *B. melitensis* and/or *B. suis*) cases (n=2) * • Rickettsial (*R. rickettsiae* and/or *R. conorii* and/or *R. typhi*) infections (n=3) *	Proven cross-reactive samples[@]: - Acute phase samples from proven Paratyphoid A, B, NTS infections -Acute phase samples collected from confirmed cases of Brucellosis, Rickettsial infections, Enterobacteriaceae infections (e.g. sepsis), and Influenza	**30[##]** -4 Paratyphoid A, 3 Paratyphoid B, 3 Paratyphoid C -1 each for the following NTS: 1. Salmonella Typhimurium 2. Salmonella Enteritidis 3. Salmonella Kentucky 4. Salmonella Eastbourne 5. Salmonella Dublin 6. Salmonella Bareilly 7. Salmonella Weltevreden 8. Salmonella Newport 9. Salmonella Infantis 10. Salmonella Agona

	4. <u>Enterobacteriaceae infection (n=3)</u> Cases of Enterobacteriaceae infection (sepsis) – including E. coli, *Klebsiella sp.*, *Citrobacter sp.* **5. <u>Influenza cases (n=2)</u>** cases of seasonal influenza virus infection (including influenza A and B)		-2 cases of Brucellosis (B. abortus and/or B. melitensis and/or B. suis) -3 cases of Rickettsial infections (R. rickettsiae and/or R. conorii and/or R. typhi)* -3 cases of Enterobacteriaceae infection (sepsis) – including E. coli, Klebsiella sp., Citrobacter sp. - 2 cases of influenza virus infections (influenza A and B)
NTS = Non Typhoidal Salmonella @ These samples should be characterized as per national guidelines/globally acceptable standards ## In case of unavailability of samples positive for a cross-reacting pathogen, the gap can be met with negative samples (negative for typhoid markers and other causes of acute febrile illness such as Dengue, malaria, chikungunya, leptospira, scrub typhus) spiked with requisite serovars. It is recommended to use well-characterized serovars from reputed and credible national/global sources or agencies.			AFI = Acute Febrile Illness NTS = Non Typhoidal Salmonella @ These samples should be characterized per national guidelines/globally acceptable standards ## In case of unavailability of samples positive for a cross-reacting pathogen, the gap can be met with negative samples (negative for typhoid markers and

* Commercially available validated panels containing genomes of these pathogens, that are accepted by accreditation agencies, may be used in case of paucity of clinical samples.		other causes of acute febrile illness as listed in the table) spiked with requisite serovars. * In case of unavailability, compensate with other samples outlined in the cross-reactivity panel

10. Statistical analysis:

Sensitivity and specificity will be calculated.

Interim analysis of data shall be conducted on completing evaluation of 25%, 50% and 75% of samples. If, at any point, the performance of the assay is found to be not satisfactory, the assay shall not be evaluated further. Evaluation fee shall be charged accordingly.

11. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples and 3 negative samples should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

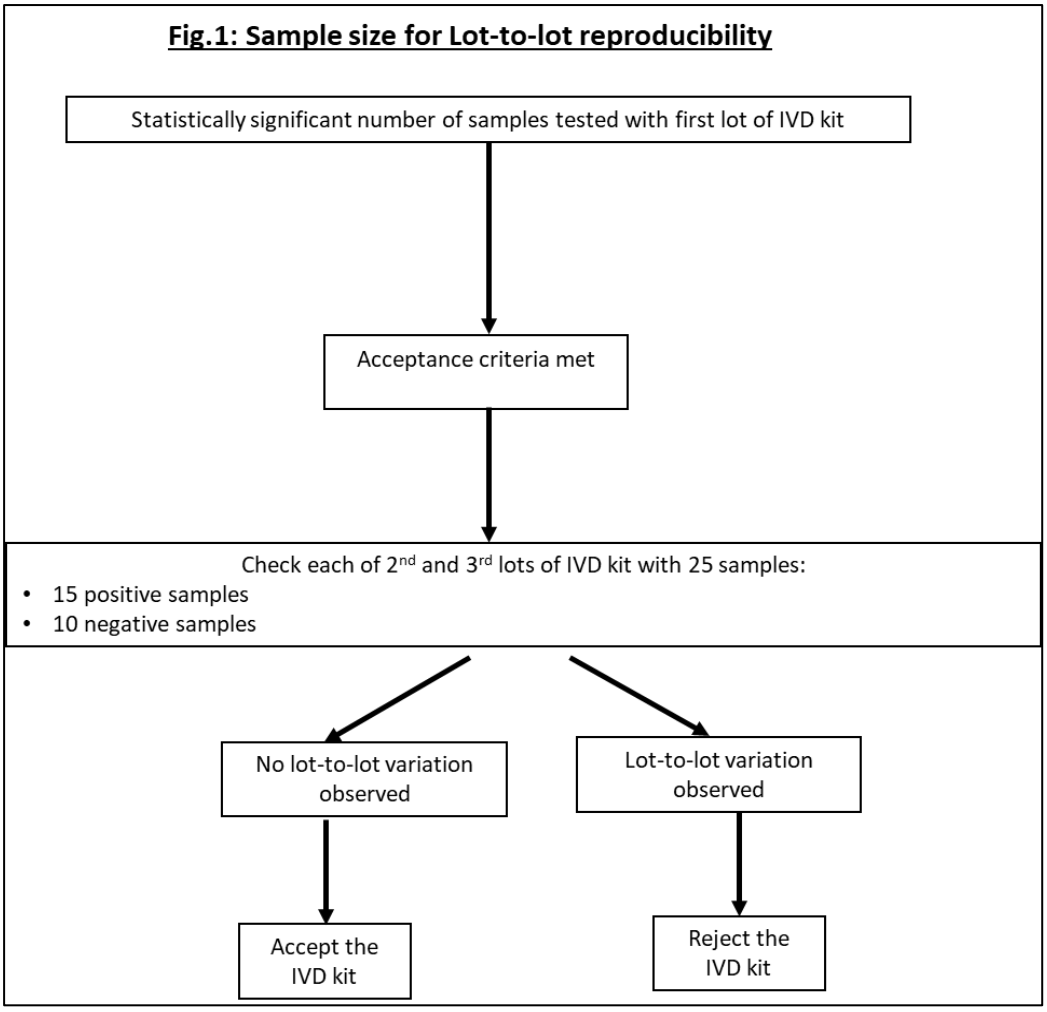
B. Reproducibility Assessment:

Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

12. Resolution of discrepant results:

True positive samples: These are samples positive by reference standard and index test.
True negative samples: These are samples negative by reference standard and index test.
False positive samples: These are samples negative by reference standard and positive by index test.
False negative samples: These are samples positive by reference standard and negative by index test.

13. Blinding of laboratory staff

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the results of the reference assay. The PI of the evaluation exercise should remain unblinded, i.e., privy to the results of the reference test. Another senior laboratory staff selected by the PI may remain unblinded for overseeing the activity and maintaining the database of results.. The data should be analyzed only by the PI of the evaluating lab. The PI should ensure data confidentiality.

14. Acceptance Criteria

Sensitivity: ≥90%

Specificity: ≥95%

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

Invalid test rate (if applicable): ≤5%

15. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

References:

1. John J, Bavdekar A, Rongsen-Chandola T, Dutta S, Gupta M, Kanungo S, et al. Burden of typhoid and paratyphoid fever in India. N Engl J Med 2023;388(16): 1491–500. https://doi.org/10.1056/NEJMoa2209449. PMID: 37075141; PMCID: PMC10116367.
2. Ministry of Health and Family Welfare-Integrated Disease Surveillance Programme. Training manual for state & district surveillance officers case definitions of diseases & syndromes under surveillance module -5. https://idsp.mohfw.gov.in/WriteReadData/OldSite/2WkDSOSept 08/Resources_files/DistrictSurvMan/Module5.pdf. [Accessed 23 November 2024].
3. Viswanathan, R., Mukhopadhyay, L., Angrup, A., Kale, P., Bhattacharyya, K., Kanungo, S., Jeromie Wesley, V. T., Shahul Hameed, S. K., Sella Senthil, M., Meshram, P., Kapil, A., Balaji, V., John, J., Ravi, V., & Gupta, N. (2025). Standard protocols for performance evaluation of typhoid fever in-vitro diagnostic assays in laboratory and field settings. *Indian journal of medical microbiology*, *55*, 100857. https://doi.org/10.1016/j.ijmmb.2025.100857
4. Mather RG, Hopkins H, Parry CM, Dittrich S. Redefining typhoid diagnosis: what would an improved test need to look like? BMJ Glob Health 2019;4(5):e001831. https://doi.org/10.1136/bmjgh-2019-001831. PMID: 31749999; PMCID: PMC6830052.
5. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

Performance Evaluation Protocol for Typhoid Antigen-based IVDs

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured diagnostic kits appropriate for use in India. Hence the following guidelines shall establish uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Typhoid antigen-based kits in the diagnosis of Typhoid fever using irreversibly de-identified leftover archived/ spiked clinical samples.

The protocol outlines performance evaluation of IVD kits that detect S Typhi antigens in blood/ serum/plasma samples.

Note: Performance evaluation with well-characterized leftover samples will be feasible once a well-performing reference molecular test and requisite clinical samples are available for use.

III. Requirements:

6. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
7. Evaluation sites/laboratories (With required equipment, and blood culture facilities)
8. Reference test kits
9. Characterised Evaluation panel
10. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.

Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.

2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through

   A. ccreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO 15189), PT provider (ISO/IEC 17043) or CDSCO approved Reference laboratory.

   B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

   - Preparation & characterization of kit evaluation panel
   - Handling of kits received for performance evaluation (Verification/Storage/Unpacking etc).
   - Testing, interpreting, recording of results & reporting
   - Data handling, data safety & confidentiality

3. **Preparation of Typhoid antigen-based IVD evaluation panel**

Well characterised sample panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of samples should be available from cases with confirmed typhoid fever status.

4. **Reference standard**:

- A composite gold standard (blood culture **AND/OR** accredited well-performing molecular test/WHO Pre-Qualified/US FDA/ ATAGI Australia/ PMDA Japan approved molecular test – e.g.: for *stag* and/or *vi* genes) should be used to characterize the samples. Use of composite gold standard is intended to increase the ability to detect a true-positive sample.
  - *Automated blood culture and ID methods should be used, and these tests should be under laboratory accreditation scope.*

It should be noted that the samples for blood culture and molecular tests may be different. It is recommended to build a biorepository of well-characterized positive and negative samples across relevant sample matrices from confirmed positive and negative cases of typhoid fever.

5. Sample size and sample panel composition*:* Sample size has been calculated assuming 95% level of significance, an absolute precision of 5%, sensitivity of ≥90%, specificity of ≥95%, and invalid test rate ≤5% (where applicable). Sample size are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples*
- *n (sp) is the minimum number of negative samples*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96)*
- *Se is the predetermined sensitivity*
- *Sp is the predetermined specificity*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

A minimum of 146 positive samples (139 samples for assays without the provision of invalid test) and a minimum of 77 negative samples (73 samples for assays without the provision of invalid test) should be used. Rounding up, a minimum of 150 positive samples and a minimum of 80 negative samples should be used for evaluation.

A. Positive samples: These samples should be acute phase samples from confirmed cases of typhoid fever positive by composite gold standard, but negative for common pathogens causing febrile illness in India (Dengue, Malaria, Scrub Typhus, Leptospira, Chikungunya).

B. Negative samples: These samples should be negative by both blood culture **AND** reference molecular test.

The sample size and sample panel composition are depicted in Table 1:

Table 1: Sample size and panel composition for performance evaluation of diagnostic antigen based kits:

A. Minimum Positive samples in the panel (N= 150)	**Minimum Number of samples needed**

Composite gold standard positive acute phase samples which are negative for common febrile illness pathogens in India[@] (dengue, malaria, scrub typhus, leptospira, chikungunya)	**150**
B. Minimum Negative samples in the panel (N=80)[#] *(samples negative by both blood culture **AND** reference molecular test)*	**Minimum Number of samples needed**
Samples from afebrile cases where blood culture and reference molecular test are negative for typhoid (e.g.: healthy blood donors) *Afebrile: absence of fever for last 14 days*	**30**
Acute phase samples from non-typhoid febrile cases– positive for dengue, malaria, scrub typhus, leptospira, chikungunya @	**20** (equal distribution for these diseases)
Cross-reactivity panel[@]: -Acute phase samples from proven Paratyphoid A, B, NTS infections -Acute phase samples from confirmed cases of Brucellosis, Rickettsial infections, Enterobacteriaceae infections (e.g. sepsis), and Influenza	**Total 30 samples (described below)** 1. **Paratyphoid samples (n=10)** • Paratyphoid A (n=4) • Paratyphoid B (n=3) • Paratyphoid C (n=3) 2. **NTS samples (n=10)** One each for the following 10 NTS[##]: • Salmonella Typhimurium • Salmonella Enteritidis • Salmonella Kentucky • Salmonella Eastbourne • Salmonella Dublin • Salmonella Bareilly • Salmonella Weltevreden • Salmonella Newport

	• Salmonella Infantis • Salmonella Agona **3.<u>Cases of Brucellosis and Rickettsial infections (n=5)</u>** • Brucellosis (B. abortus and/or *B. melitensis* and/or *B. suis*) cases (n=2) * • Rickettsial (*R. rickettsiae* and/or *R. conorii* and/or *R. typhi*) infections (n=3) * **4. <u>Enterobacteriaceae infection (n=3)</u>** Cases of Enterobacteriaceae infection (sepsis) – including E. coli, *Klebsiella sp.*, *Citrobacter sp.* **5. <u>Influenza cases (n=2)</u>** cases of seasonal influenza virus infection (including influenza A and B)

NTS = Non Typhoidal Salmonella

These samples should be characterized as per national guidelines/globally acceptable standards

In case of unavailability of samples positive for a cross-reacting pathogen, the gap can be met with negative samples (negative for typhoid markers and other causes of acute febrile illness such as Dengue, malaria, chikungunya, leptospira, scrub typhus) spiked with requisite serovars. It is recommended to use well-characterized serovars from reputed and credible national/global sources or agencies.

* Commercially available validated panels containing genomes of these pathogens, that are accepted by accreditation agencies, may be used in case of paucity of clinical samples.

6. Evaluation method:

It is recommended to run the reference molecular test and the index test in parallel.

7. Interpretation of results:

Reference test and index test results will be interpreted as per kit IFU/SoP.

8. Resolution of discrepant results:

True positive samples: These are samples positive by reference test and index test.

True negative samples: These are samples negative by reference test and index test.

False positive samples: These are samples negative by reference test and positive by index test.

False negative samples: These are samples positive by reference test and negative by index test.

9. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples and 3 negative samples should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

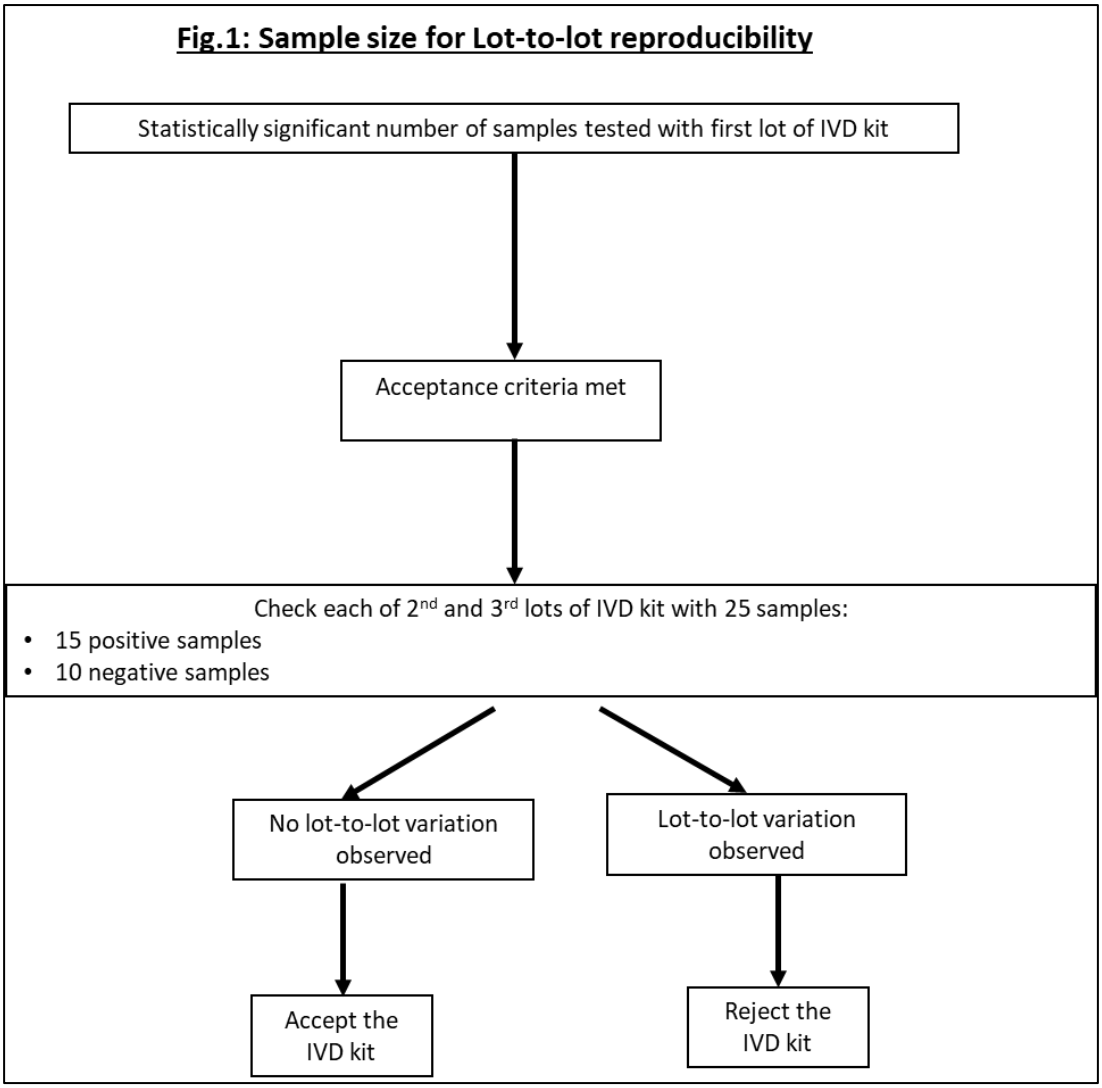
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Fig.1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if

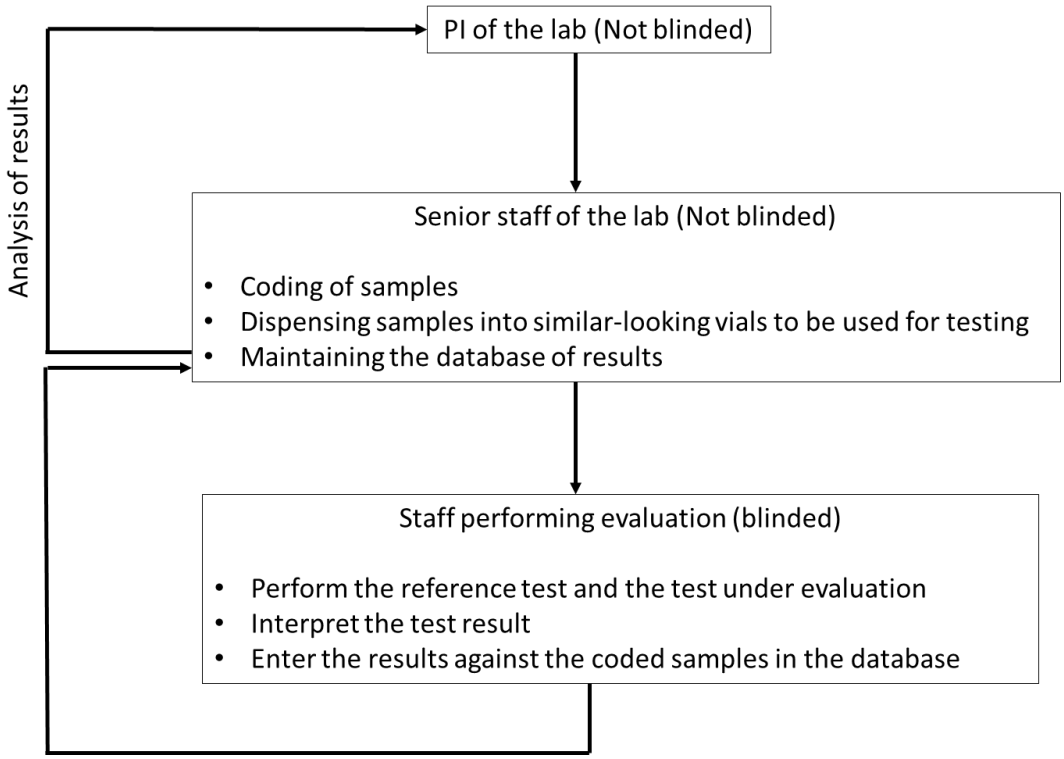
applicable). 3 positive samples and 3 negative samples should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

10. Blinding of laboratory staff

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab, and the PI should ensure confidentiality of data. Refer to Fig. 1.

Fig.1: Blinding in evaluation exercise



11. Acceptance Criteria

Sensitivity: ≥90%

Specificity: ≥95%

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

Invalid test rate(if applicable): ≤5%

12. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

References:

1. John J, Bavdekar A, Rongsen-Chandola T, Dutta S, Gupta M, Kanungo S, et al. Burden of typhoid and paratyphoid fever in India. N Engl J Med 2023;388(16): 1491–500. https://doi.org/10.1056/NEJMoa2209449. PMID: 37075141; PMCID: PMC10116367.
2. Ministry of Health and Family Welfare-Integrated Disease Surveillance Programme. Training manual for state & district surveillance officers case definitions of diseases & syndromes under surveillance module -5. https://idsp.mohfw.gov.in/WriteReadData/OldSite/2WkDSOSept 08/Resources_files/DistrictSurvMan/Module5.pdf. [Accessed 23 November 2024].
3. Viswanathan, R., Mukhopadhyay, L., Angrup, A., Kale, P., Bhattacharyya, K., Kanungo, S., Jeromie Wesley, V. T., Shahul Hameed, S. K., Sella Senthil, M., Meshram, P., Kapil, A., Balaji, V., John, J., Ravi, V., & Gupta, N. (2025). Standard protocols for performance evaluation of typhoid fever in-vitro diagnostic assays in laboratory and field settings. *Indian journal of medical microbiology*, *55*, 100857. https://doi.org/10.1016/j.ijmmb.2025.100857
4. Mather RG, Hopkins H, Parry CM, Dittrich S. Redefining typhoid diagnosis: what would an improved test need to look like? BMJ Glob Health 2019;4(5):e001831. https://doi.org/10.1136/bmjgh-2019-001831. PMID: 31749999; PMCID: PMC6830052.
5. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

Field Evaluation Protocol for Typhoid Antibody-based IVDs

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured Diagnostic kits appropriate for use in India. Hence the following guidelines shall establish the uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Typhoid antibody-based kits in the diagnosis of Typhoid fever in individuals with unknown disease status.

The protocol outlines field evaluation of IVD kits that detect S Typhi antibodies in blood/ serum/plasma samples.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment)
3. Reference test kits
4. Laboratory supplies

IV. Ethical approval:

The study will be initiated after obtaining approval from the institutional human ethics committee.

V. Procedure:

1. Study design/type: Cross-sectional study

2. Preparation of Evaluation sites/laboratories:

Identified IVD kit evaluation laboratories should establish their proficiency through

A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IEC 17025), Medical Lab (ISO/IEC 15189), PT provider (ISO/IEC 17043) or CDSCO approved Reference laboratory.

B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

- Preparation & characterization of kit evaluation panel
- Handling of Typhoid antibody-based IVD received for performance evaluation (Verification/Storage/Unpacking etc).
- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Sample size for performance evaluation:

Sample size has been calculated assuming 95% level of significance, an absolute precision of 5%, sensitivity ≥80%, specificity ≥90%, and invalid test rate ≤5% (where applicable). Sample size is calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})xP}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})xP}$$

- *n (se) is the minimum number of positive samples*
- *n (sp) is the minimum number of negative samples*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96)*
- *Se is the predetermined sensitivity*
- *Sp is the predetermined specificity*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*
- *P is the prevalence of the disease*

A minimum of 6500 cases with acute febrile illness satisfying the case definition need to be enrolled in the study to achieve statistically significant minimum number of positive samples.

Explanation: A minimum of 259 positive cases (246 cases for assays without the provision of invalid test) and a minimum of 146 negative cases (139 cases for assays without the provision of invalid test) should be enrolled as per the aforementioned formulae. Rounding up, a minimum of 260 positive cases and a minimum of 150 negative cases should be enrolled for evaluation. Since approximately 4% of acute fever cases are positive for typhoid through blood culture (John J, 2023), minimum 6500 cases satisfying the case definition need to be enrolled to achieve the desired target of minimum 260 positive cases. This sample size is adequate for the desired number of negative cases that need to be tested in the evaluation.

Recruitment of cases shall be halted once desired number of positive and negative samples are reached.

Note: Typhoid being a seasonal disease, attempts should be made to enrol more cases during seasons of high transmission. For multi-centric studies, clinicians at each participating site should be sensitized on case enrolment.

4. Inclusion criteria:

The case definition from India's Integrated Disease Surveillance Programme (IDSP) has been adapted to increase the probability of enrolment of typhoid blood culture positive, and ensure availability of adequate sample volumes (Ministry of Health & Family Welfare, 2024)**.** Briefly, an individual aged 2-65 years from OPD/IPD settings will be included if they satisfy both the major and any one of the minor criteria mentioned below:

Major criteria: Documented fever of >38°C **AND** Duration of fever 4-10 days (preferably ≥3 consecutive days of fever)

Minor criteria:
- Toxic look
- Coated tongue
- Relative bradycardia
- Splenomegaly

5. Exclusion criteria:

- Age <2 years or ≥65 years
- Pregnancy
- Chronic diseases
- Known or identified focus of infection

6. Reference standard and sample characterization:

a. Sample collection: There is no predicate IVD kit/device available for use as diagnostic typhoid antibody reference standard, and cross-reactivity and anamnestic responses significantly impact the specificity of these assays. Therefore, it is important to ensure characterization of the sample panel to be used for evaluation. The study should enrol acutely febrile cases satisfying the case definition, and collect blood samples for different tests for identifying true positive typhoid cases at various time points as outlined in Table 1.

Table 1: Sample type and Timelines for Collection:

Age of participants	**Sample/Timing**	**Antibody detecting IVDs**
Adults	Sample type and timing of collection	1.*Blood in 2 automated blood culture bottles (10-20 ml blood in each bottle) from 2 different sampling sites* *Timing:* on case enrolment (acute febrile phase) 2. *3-4 ml blood in EDTA vacutainer (for reference molecular test)* *Timing:* on case enrolment (acute febrile phase) 3. *3-4 ml blood in plain vacutainer (for accredited Widal test-baseline antibody titers)* *Timing:* on case enrolment (acute febrile phase)

		4. *Follow up Sample: 3-4 ml blood in plain vacutainer (for accredited Widal test and index test)* *Timing:* 3-4 weeks after onset of illness (which should also in accordance with the duration of illness mentioned by the manufacturer)
Children	Sample type and timing of collection	1.Blood in 1 automated blood culture bottle (3-5 ml of blood for children aged 2-5 years and 5-10 ml of blood for children aged >5-13 years) from 1 sampling site *Timing:* on case enrolment (febrile phase, in accordance with the duration of illness mentioned by the manufacturer) 2. 3-4 ml blood in EDTA vacutainer (for reference molecular test) *Timing:* on case enrolment (febrile phase, in accordance with the duration of illness mentioned by the manufacturer) 3. 3-4 ml blood in plain vacutainer *(for accredited Widal test-baseline antibody titers)* *Timing:* on case enrolment (febrile phase, in accordance with the duration of illness mentioned by the manufacturer) 4. Follow up Sample: 3-4 ml blood in plain vacutainer (for accredited Widal test and index test) *Timing:* 3-4 weeks after onset of illness (if needed, also in accordance with the duration of illness mentioned by the manufacturer)

If an antibody-based IVD kit claims to detect anti-typhoid antibodies in samples collected earlier in the course of the disease (less than 3-4 weeks), samples should be available from these time points as well for evaluation. If these samples are not available, the performance evaluation report should clearly indicate the timing of follow up sample collection.

b. Confirmation of typhoid fever status and obtaining typhoid antibody-positive samples from convalescent cases:

i. *On enrolment:* Typhoid fever status should be confirmed for all cases on enrolment through a composite reference standard of blood culture **AND/OR** standardized well-performing rigorously validated molecular test (under laboratory accreditation scope/WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan

approved assays). Use of composite gold standard is intended to increase the ability to detect a true-positive case.

- *Blood culture alone should be used for determining typhoid fever status till the time a well-performing reference molecular test (as outlined in this section) is available with the evaluating labs.*
- *Automated blood culture and ID methods should be used, and these tests should be under laboratory accreditation scope.*
- *Note: WHO Pre-Qualified/US FDA/ATAGI Australia/ PMDA Japan approved Typhoid molecular kit may be used as reference molecular test as and when these kits become available.*
- An accredited Widal test (or WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test as and when they become available) should be performed on the accompanying serum sample collected on enrolment to establish baseline anti-typhoid antibody titer.

ii. *Follow-up sample*: A follow-up serum sample should be collected 3-4 weeks after the onset of illness to confirm the presence of anti-typhoid antibodies through Widal test.

Positive cases: Confirmed typhoid fever cases should demonstrate 4-fold rise in Widal TO/TH titers (>320) in samples collected after 3-4 weeks; this sample panel should comprise anti-typhoid antibody positive samples and should be used for evaluation of the index test.

Negative cases: Sera samples collected 3-4 weeks after onset of illness from cases without typhoid fever and not showing 4-fold rise in Widal TO/TH titers (>320), should be considered negative samples, and should be used for evaluation of the index test.

Note: Blood culture/molecular test negative typhoid cases:

For a blood culture **AND** reference molecular test negative (as per the conditions outlined in Section 6) but baseline Widal positive (TH titer > 320) case to be considered a true positive case of typhoid, the case should demonstrate at least 4-fold rise in antibody titers when follow-up samples are collected prospectively 3-4 weeks after onset of illness, in the absence of another confirmed cause of fever that may lead to anamnestic O antigen-related responses. Such samples may also be included in the positive sample panel for performance evaluation.

Note:

- All Widal tests performed for sample characterization as per Section 6 should be accredited as per ISO 15189 or ISO/IEC 17025 or ISO/IEC 17043.
- WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test may be used as and when they become available.

7. Study implementation:

The individuals satisfying the case definition will be recruited, and samples will be collected and tested as per the requirements of the study.

The requisite samples collected 3-4 weeks after onset of illness as defined in Section 6 should be tested by accredited Widal test (or WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test as and when they become available) and the index test simultaneously.

8. Interpretation of results:

All test results (blood culture, reference molecular test, Widal test, index test) will be interpreted as per respective kit IFU/SoP.

9. Resolution of discrepant results:

- **True Positive samples**:

Rigorously well-characterized sera samples collected 3-4 weeks after onset of illness from confirmed typhoid cases (**as per Section 6**) and showing TO/TH titers >320 with accredited Widal test (or WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test as and when they become available), and positive results with the index test, will be considered as true positive sample.

For semi-quantitative antibody-based tests, results should closely mimic Widal test results.

- **True Negative samples**:

Rigorously well-characterized sera samples collected 3-4 weeks after onset of illness from confirmed non-typhoid cases (**as per Section 6**) and testing non-reactive (or) for anti-typhoid antibodies with accredited Widal test (or WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test as and when they become available), as well as the index test, will be considered as true negative sample.

- **False positive samples:**

Rigorously well-characterized sera samples collected 3-4 weeks after onset of illness from confirmed non-typhoid cases (**as per Section 6**) and testing non-reactive for anti-typhoid antibodies with accredited Widal test (or WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test as and when they become available), but reactive for anti-typhoid antibodies with the index test, will be considered as false positive samples.

- **False negative samples:**

Rigorously well-characterized sera samples collected 3-4 weeks after onset of illness from confirmed typhoid cases (**as per Section 6**) and showing TO/TH titers >320 with accredited Widal test (or WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test as and when they become available), but non-reactive for anti-typhoid antibodies with the index test, will be considered false negative samples.

10. Cross-reactivity: Cross-reactivity should be reported separately using the samples mentioned below:

Proven cross-reactive samples: - Sera collected 2-4 weeks after onset of illness in proven Paratyphoid A, B, and NTS infections (blood culture/PCR confirmed cases) @ -Sera collected 2-4 weeks after onset of illness from confirmed cases of Brucellosis, Rickettsial infections, Enterobacteriaceae infections (e.g. sepsis), and Influenza	**Recommended: 30 samples** **Minimum: 25 samples (described below)** **1. Paratyphoid samples (n=10)** • Paratyphoid A (n=4) • Paratyphoid B (n=3) • Paratyphoid C (n=3) 2. **NTS samples (n=10)** One each for the following 10 NTS##: • Salmonella Typhimurium • Salmonella Enteritidis • Salmonella Kentucky • Salmonella Eastbourne • Salmonella Dublin • Salmonella Bareilly • Salmonella Weltevreden • Salmonella Newport • Salmonella Infantis • Salmonella Agona **3.Cases of Brucellosis and Rickettsial infections (n=5)** • Brucellosis (B. abortus and/or *B. melitensis* and/or *B. suis*) cases (n=2) * • Rickettsial (*R. rickettsiae* and/or *R. conorii* and/or *R. typhi*) infections (n=3) * ‡ **4. Enterobacteriaceae infection (n=3)** Cases of Enterobacteriaceae infection (sepsis) – including E. coli, *Klebsiella sp.*, *Citrobacter sp.* **5. Influenza cases (n=2)** cases of seasonal influenza virus infection (including influenza A and B)

NTS = Non Typhoidal Salmonella

In case of unavailability of samples positive for a cross-reacting pathogen, attempt should be made to meet the gap with samples from clinical cases of infection with other serovars. If deficit persists, the gap should be met with samples from Paratyphoid fever cases. If no sample positive

from confirmed cases of these infections is available, test 5 more paratyphoid fever positive samples (in addition to the 10 samples outlined in Point No. 1); this will bring down the size of the cross-reactivity panel to 25 samples. Commercially available serology panels for these pathogens, that are accepted by accreditation agencies, may also be used.

* Commercially available serology panels for these pathogens, that are accepted by accreditation agencies, may also be used in case of paucity of clinical samples.

‡ In case of unavailability/paucity of Rickettsial serology samples, the gap should be met with scrub typhus IgM antibody positive samples.

11. Statistical analysis:

Sensitivity and specificity will be calculated.

Interim analysis of data shall be conducted on completing evaluation of 25%, 50% and 75% of samples. If, at any point, the performance of the assay is found to be not satisfactory, the assay shall not be evaluated further. Evaluation fee shall be charged accordingly.

12. Repeatability and Reproducibility Assessment:

A. Repeatability Assessment

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples and 3 negative samples should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

Reproducibility testing should include the following:

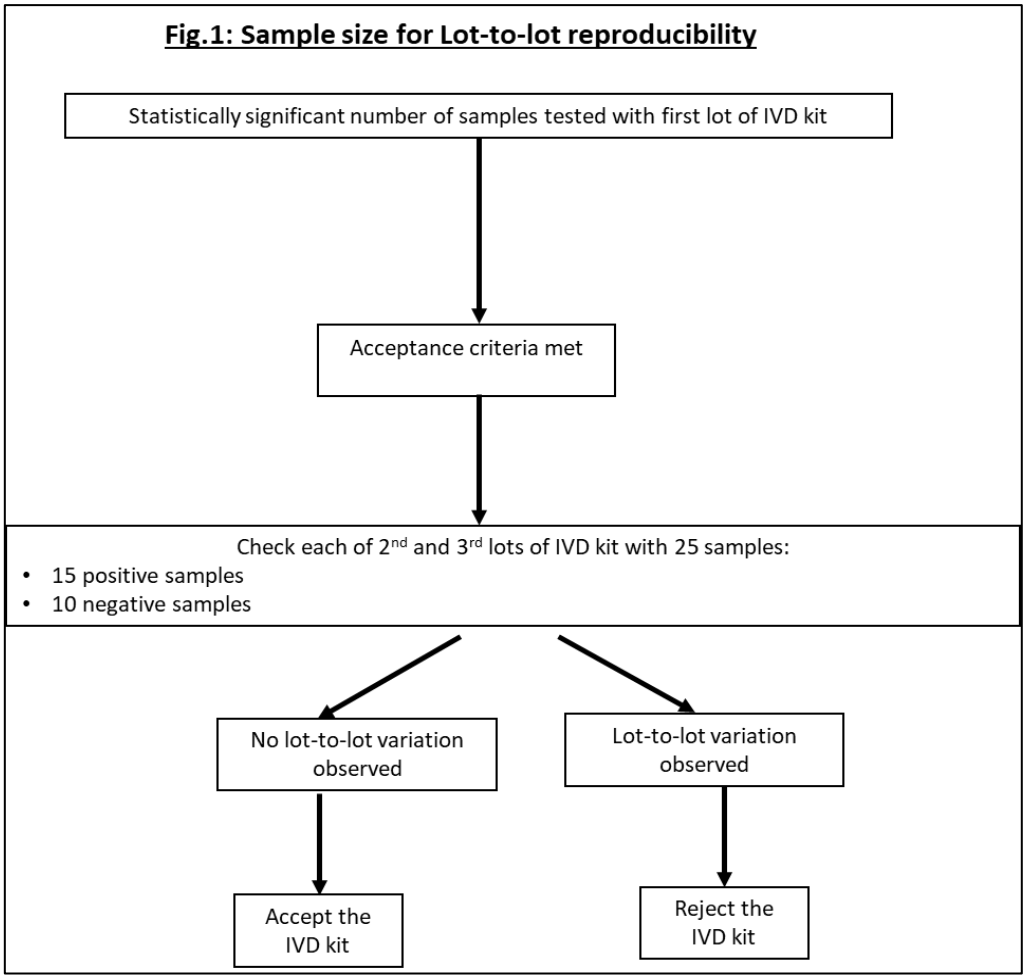
a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.

- Second lot of the kit: should be tested on 25 samples (15 positive samples and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Fig.1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if

applicable). 3 positive samples and 3 negative samples should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

13. Blinding of laboratory staff

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the results of the reference assay. The PI of the evaluation exercise should remain unblinded, i.e., privy to the results of the reference test. Another senior laboratory staff selected by the PI may remain unblinded for overseeing the activity and maintaining the database of results. The data should be analyzed only by the PI of the evaluating lab. Data confidentiality should be ensured by the PI.

14. Acceptance Criteria

Sensitivity: ≥80%

Specificity: ≥90%

Cross-reactivity with other pathogens listed in the negative sample panel: Minimal

Invalid test rate (if applicable): ≤5%

15. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

References:

1. John J, Bavdekar A, Rongsen-Chandola T, Dutta S, Gupta M, Kanungo S, et al. Burden of typhoid and paratyphoid fever in India. N Engl J Med 2023;388(16): 1491–500. https://doi.org/10.1056/NEJMoa2209449. PMID: 37075141; PMCID: PMC10116367.
2. Ministry of Health and Family Welfare-Integrated Disease Surveillance Programme. Training manual for state & district surveillance officers case definitions of diseases & syndromes under surveillance module -5. https://idsp.mohfw.gov.in/WriteReadData/OldSite/2WkDSOSept 08/Resources_files/DistrictSurvMan/Module5.pdf. [Accessed 23 November 2024].

3. Viswanathan, R., Mukhopadhyay, L., Angrup, A., Kale, P., Bhattacharyya, K., Kanungo, S., Jeromie Wesley, V. T., Shahul Hameed, S. K., Sella Senthil, M., Meshram, P., Kapil, A., Balaji, V., John, J., Ravi, V., & Gupta, N. (2025). Standard protocols for performance evaluation of typhoid fever in-vitro diagnostic assays in laboratory and field settings. *Indian journal of medical microbiology*, *55*, 100857. https://doi.org/10.1016/j.ijmmb.2025.100857
4. Mather RG, Hopkins H, Parry CM, Dittrich S. Redefining typhoid diagnosis: what would an improved test need to look like? BMJ Glob Health 2019;4(5):e001831. https://doi.org/10.1136/bmjgh-2019-001831. PMID: 31749999; PMCID: PMC6830052.
5. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

Performance Evaluation Protocol for Typhoid Antibody-based IVDs

I. Background:

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured diagnostic kits appropriate for use in India. Hence the following guidelines shall establish uniformity in performance evaluation of in-vitro diagnostic kits (IVD). The performance evaluation is to independently verify the manufacturer's claim regarding IVD performance.

II. Purpose:

To evaluate the performance characteristics of Typhoid antibody-based kits in the diagnosis of Typhoid fever using irreversibly de-identified leftover archived clinical samples.

The protocol outlines performance evaluation of IVD kits that detect S Typhi antibodies in blood/ serum/plasma samples.

Note: Performance evaluation with leftover clinical samples will be feasible only when well-characterized samples, as outlined in the document, are available for use.

III. Requirements:

1. Supply of kits under evaluation (Along with batch/lot No. Expiry & required details). If the kit to be evaluated works in a closed system format, the manufacturer needs to supply the required equipment.
2. Evaluation sites/laboratories (With required equipment, and blood culture facilities)
3. Reference test kits
4. Characterised Evaluation panel
5. Laboratory supplies

IV. Ethical approvals:

Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR's Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.
Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

V. Procedure:

1. **Study design/type**: Diagnostic accuracy study using irreversibly de-identified leftover clinical/spiked samples.
2. **Preparation of Evaluation sites/laboratories:**
   Identified IVD kit evaluation laboratories should establish their proficiency through

A. Accreditation for at least one Quality management system (accreditation for Testing Lab / Calibration Lab (ISO/IES 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory.

B. Staff training: All the staff involved in IVD kit evaluation should undergo hands on training and competency testing on following

- Preparation & characterization of kit evaluation panel
- Handling of kits received for performance evaluation (Verification/Storage/Unpacking etc).
- Testing, interpreting, recording of results & reporting
- Data handling, data safety & confidentiality

3. Preparation of Typhoid antibody-based IVD evaluation panel

Well characterised sample panel is a critical requirement for performance evaluation of IVD kits. Hence statistically significant number of well-characterized samples should be available from confirmed Typhoid fever cases.

4. Sample size and sample panel composition: Sample size has been calculated assuming 95% level of significance, an absolute precision of 5%, sensitivity of ≥80%, specificity of ≥90%, and invalid test rate ≤5% (where applicable). Sample size are calculated using the formulae:

$$n_{se} \geq \frac{Z^2 \text{ x } S_e\ (1 - S_e)}{d^2 \text{ x } (1 - \text{IR})}$$

$$n_{sp} \geq \frac{Z^2 \text{ x } S_p\ (1 - S_p)}{d^2 \text{ x } (1 - \text{IR})}$$

- *n (se) is the minimum number of positive samples*
- *n (sp) is the minimum number of negative samples*
- *Z^2 is the critical value from the standard normal distribution corresponding to the desired confidence level (95% CI corresponds to Z^2 =1.96)*
- *Se is the predetermined sensitivity*
- *Sp is the predetermined specificity*
- *d is the predetermined marginal error (5%)*
- *IR is the invalid test rate*

A minimum of 259 positive samples (246 samples for assays without the provision of invalid test) and a minimum of 146 negative samples (139 samples for assays without the provision of invalid test). Rounding up, a minimum of 260 positive samples and a minimum of 150 negative samples should be used for evaluation.

5. Sample panel characterization and Reference standard:

True positive and negative samples should be characterized as per the "Field Evaluation Protocol for Typhoid Antibody-based IVDs" (Sections 6 and 8 of the document).

The sample size and sample panel composition are depicted in Table 1:

Table 1: Sample size and panel composition for analytical evaluation of diagnostic antibody-based assays:

A. Positive samples in the panel (N= 260)	**Minimum Number of samples needed**
True positive samples as described in "Field Evaluation Protocol for Typhoid Antibody-based IVDs" (Sections 6 and 8 of the document).	**260**
B. Negative samples in the panel (minimum N=150) *True negative samples should satisfy the conditions described in "Field Evaluation Protocol for Typhoid Antibody-based IVDs" (Sections 6 and 8 of the document).*	**Minimum Number of samples needed**
Sera from confirmed afebrile non-typhoid cases **AND** negative for anti-typhoid antibodies (e.g.: healthy blood donors)	**80**
Sera collected 2-4 weeks after onset of illness in acutely febrile confirmed non-typhoid cases (samples positive for dengue, malaria, scrub typhus, leptospira, chikungunya)[@]	**40** Dengue: 14 (desirable to have representation from all 4 serotypes) Malaria: 5 Scrub typhus: 5 Leptospira: 8 Chikungunya: 8
Proven cross-reactive samples: - Sera collected 2-4 weeks after onset of illness in proven Paratyphoid A, B, and NTS infections (blood culture/PCR confirmed cases) [@] -Sera collected 2-4 weeks after onset of illness from confirmed cases of Brucellosis, Rickettsial infections, Enterobacteriaceae infections (e.g. sepsis), and Influenza	**Recommended: 35 samples** **Minimum: 30 samples (described below)** **2. Paratyphoid samples (n=10)** • Paratyphoid A (n=4) • Paratyphoid B (n=3) • Paratyphoid C (n=3) 2. **NTS samples (n=10)** One each for the following 10 NTS[##]: • Salmonella Typhimurium • Salmonella Enteritidis • Salmonella Kentucky • Salmonella Eastbourne • Salmonella Dublin • Salmonella Bareilly • Salmonella Weltevreden • Salmonella Newport • Salmonella Infantis • Salmonella Agona **3.Cases of Brucellosis and Rickettsial infections (n=8)** • Brucellosis (B. abortus and/or *B.*

	melitensis and/or *B. suis*) cases (n=3) * • Rickettsial (*R. rickettsiae* and/or *R. conorii* and/or *R. typhi*) infections (n=5) *‡ **4. <u>Enterobacteriaceae infection (n=5)</u>** Cases of Enterobacteriaceae infection (sepsis) – including E. coli, *Klebsiella sp., Citrobacter sp.* **6. <u>Influenza cases (n=2)</u>** cases of seasonal influenza virus infection (including influenza A and B)

NTS = Non Typhoidal Salmonella

@ These samples should be characterized as per national guidelines/globally acceptable standards.

In case of unavailability of samples positive for a cross-reacting pathogen, attempt should be made to meet the gap with samples from clinical cases of infection with other serovars. If deficit persists, the gap should be met with samples from Paratyphoid fever cases. If no sample positive from confirmed cases of these infections is available, test 5 more paratyphoid fever positive samples (in addition to the 10 samples outlined in Point No. 1); this will bring down the size of the cross-reactivity panel to 30 samples. Commercially available serology panels for these pathogens, that are accepted by accreditation agencies, may also be used.

* Commercially available serology panels for these pathogens, that are accepted by accreditation agencies, may also be used in case of paucity of clinical samples.

‡ In case of unavailability/paucity of Rickettsial serology samples, the gap should be met with scrub typhus IgM antibody positive samples.

Well-characterized true positive and true negative samples as outlined in Table 1 should be subjected to an accredited Widal test (or WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test as and when they become available) to ensure sample integrity at the time of evaluation.

7. Evaluation method:

The accredited Widal test (or WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test as and when they become available) and the index test should be run simultaneously on the sample panel.

8. Interpretation of results:

All test results will be interpreted as per respective kit IFU/SoP.

9. Resolution of discrepant results:

- **True Positive samples**:

Rigorously well-characterized sera samples collected 3-4 weeks after onset of illness from confirmed typhoid cases (see Note below) and showing TO/TH titers >320 with accredited Widal test (or WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test as and when they become available), and positive results with the index test, will be considered as true positive sample.

For semi-quantitative antibody-based tests, results should closely mimic Widal test results.

- **True Negative samples**:

Rigorously well-characterized sera samples collected 3-4 weeks after onset of illness from confirmed non-typhoid cases (see Note below) and testing non-reactive for anti-typhoid antibodies with accredited Widal test (or WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test as and when they become available), as well as the index test, will be considered as true negative sample.

- **False positive samples:**

Rigorously well-characterized sera samples collected 3-4 weeks after onset of illness from confirmed non-typhoid cases (see Note below) and testing non-reactive for anti-typhoid antibodies with accredited Widal test (or WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test as and when they become available), but reactive for anti-typhoid antibodies with the index test, will be considered as false positive samples.

- **False negative samples:**

Rigorously well-characterized sera samples collected 3-4 weeks after onset of illness from confirmed typhoid cases (see Note below) and showing TO/TH titers >320 with accredited Widal test (or WHO Pre-Qualified/US FDA/ATAGI Australia/PMDA Japan approved Widal test as and when they become available), but non-reactive for anti-typhoid antibodies with the index test, will be considered false negative samples.
Note: The sample panel should be characterized as per the "Field Evaluation Protocol for Typhoid Antibody-based IVDs" (Sections 6 and 8 of the document).

10. Repeatability and Reproducibility Assessment:

A. <u>Repeatability Assessment</u>

This should be done to assess the repeatability of the detection of target using the kit under evaluation.

3 positive samples and 3 negative samples should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

B. Reproducibility Assessment:

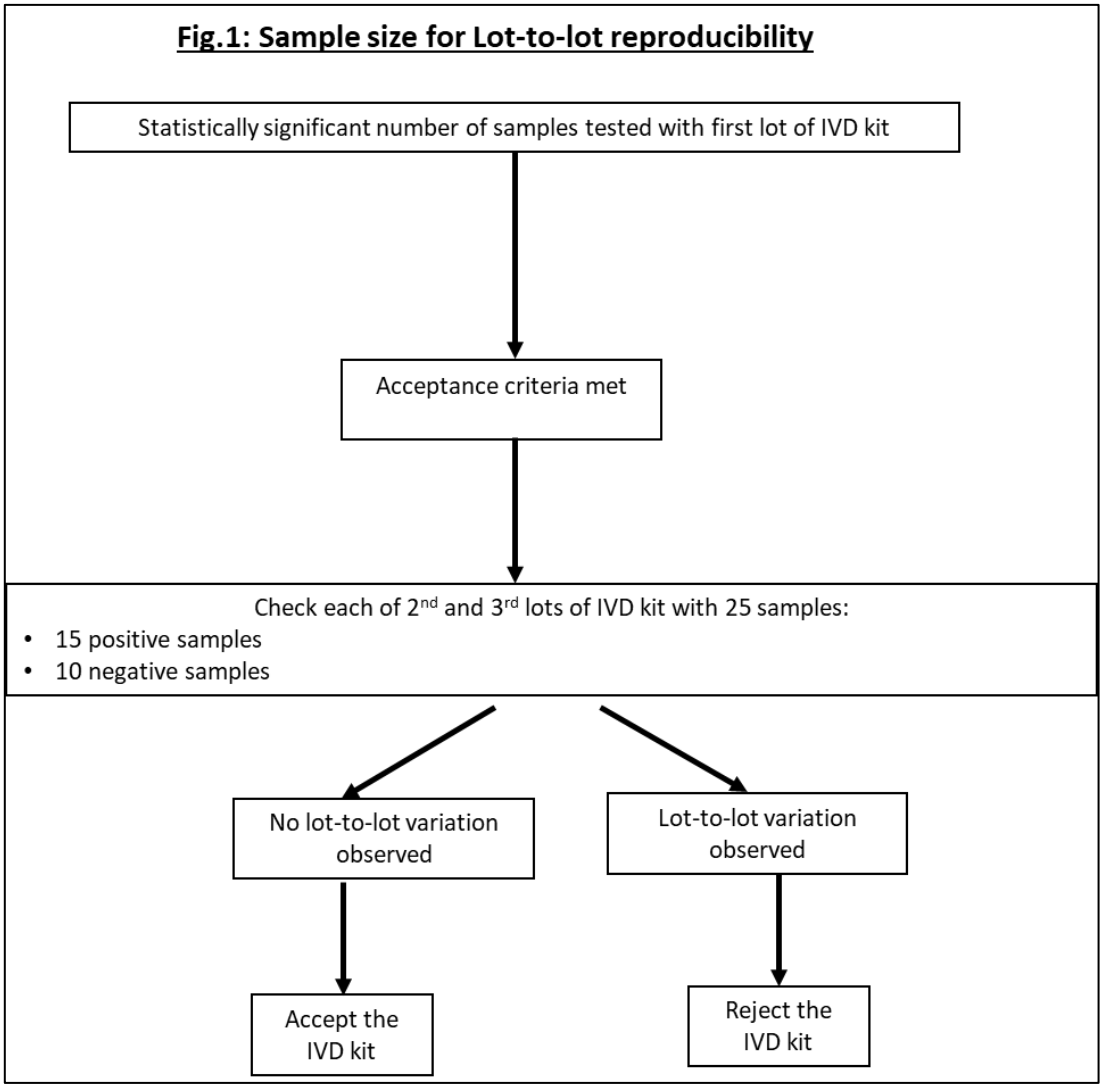
Reproducibility testing should include the following:

a. Lot-to-lot reproducibility

Three lots of an IVD kit shall be evaluated. Sample size for lot-to-lot reproducibility should be as follows:

- First lot of the kit: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the kit: should be tested on 25 samples (15 positive samples and 10 negative samples).
- Third lot of the kit: should be tested on 25 samples (15 positive samples and 10 negative samples).
- There should be no lot-to-lot variability (qualitative). Refer Figure 1.

Fig.1: Sample size for Lot-to-lot reproducibility



b. Inter-Operator variability: Testing should be conducted by two different operators, keeping all other testing parameters undisturbed. Within-run and between-run imprecision (if applicable) should be measured. 3 positive samples and 3 negative samples should be tested 5 times in independent runs.

Concordance should be 100% based on positive and negative test result (qualitative).

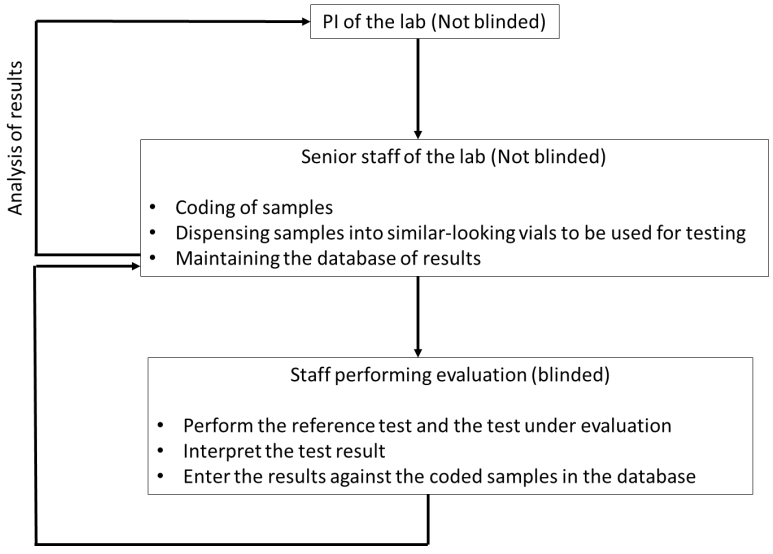
c. Day-to-day variability: Testing should be performed on at least two non-consecutive days. 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

d. Machine-to-machine variability: It is desirable (not mandatory) to evaluate the IVD kit using two different manufacturer recommended platforms (if applicable). 3 positive samples and 3 negative samples should be tested 5 times in independent runs.
Concordance should be 100% based on positive and negative test result (qualitative).

10. Blinding of laboratory staff:

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab. The PI should maintain confidentiality of the data. Refer to Fig. 1.

Fig.1: Blinding in evaluation exercise



11. Acceptance Criteria

Sensitivity: ≥80%
Specificity: ≥90%
Cross-reactivity with other pathogens listed in the negative sample panel: Minimal
Invalid test rate (if applicable): ≤5%

12. Publication Rights:

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.
Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.
Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

References:

1. John J, Bavdekar A, Rongsen-Chandola T, Dutta S, Gupta M, Kanungo S, et al. Burden of typhoid and paratyphoid fever in India. N Engl J Med 2023;388(16): 1491–500. https://doi.org/10.1056/NEJMoa2209449. PMID: 37075141; PMCID: PMC10116367.
2. Ministry of Health and Family Welfare-Integrated Disease Surveillance Programme. Training manual for state & district surveillance officers case definitions of diseases & syndromes under surveillance module -5. https://idsp.mohfw.gov.in/WriteReadData/OldSite/2WkDSOSept 08/Resources_files/DistrictSurvMan/Module5.pdf. [Accessed 23 November 2024].
3. Viswanathan, R., Mukhopadhyay, L., Angrup, A., Kale, P., Bhattacharyya, K., Kanungo, S., Jeromie Wesley, V. T., Shahul Hameed, S. K., Sella Senthil, M., Meshram, P., Kapil, A., Balaji, V., John, J., Ravi, V., & Gupta, N. (2025). Standard protocols for performance evaluation of typhoid fever in-vitro diagnostic assays in laboratory and field settings. *Indian journal of medical microbiology*, *55*, 100857. https://doi.org/10.1016/j.ijmmb.2025.100857
4. Mather RG, Hopkins H, Parry CM, Dittrich S. Redefining typhoid diagnosis: what would an improved test need to look like? BMJ Glob Health 2019;4(5):e001831. https://doi.org/10.1136/bmjgh-2019-001831. PMID: 31749999; PMCID: PMC6830052.
5. Monti, C. B., Ambrogi, F., & Sardanelli, F. (2024). Sample size calculation for data reliability and diagnostic performance: a go-to review. *European radiology experimental*, *8*(1), 79. https://doi.org/10.1186/s41747-024-00474-w

Performance evaluation report format for Typhoid IVDs

Name of the product (Brand /generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No / Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number: Issue date: Valid Up to:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: clinical/spiked samples) ***If the index test claims to detect the disease during a period of illness for which samples are not available, clearly mention the disease period covered by the sample panel used for evaluation.***	

	Negative samples (provide details samples)	

Results

		Reference assay ……..…… (name)		
		Positive	Negative	Total
Name of index Typhoid IVD kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

- Cross-reactivity:
- Invalid test rate:

FINAL CONCLUSION

Performance: Satisfactory / Not satisfactory

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ….. sample. Results should not be extrapolated to other sample types.)

Disclaimers

1. This validation process does not approve / disapprove the kit design

2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for …….. Kit (Lot No……) manufactured by …… (supplied by …….)

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge ……
Seal…….

Information on Operational and Test Performance Characteristics Required from Manufacturers

The manufacturer should provide the following details about the IVD:

1. Instructions for Use
2. Scope of the IVD:
3. Intended Use Statement
4. Principle of the assay
5. Intended testing population
6. Intended user (laboratory professional and/or health care worker at point-of-care)
7. Detailed test protocol
8. Lot/batch No.
9. Date of manufacture
10. Date of Expiry
11. Information on operational Characteristics
    - i. Configuration of the kit/device
    - ii. Requirement of any additional equipment, device
    - iii. Requirement of any additional reagents
    - iv. Operation conditions
    - v. Storage and stability before and after opening
    - vi. Internal control provided or not
    - vii. Quality control and batch testing data
    - viii. Biosafety aspects- waste disposal requirements

10. Information on Test Performance Characteristics
    - i. Type of sample-serum/plasma/whole blood/other specimen (specify)
    - ii. Volume of sample
    - iii. Sample rejection criteria (if any)
    - iv. Any additional sample processing required
    - v. Any additional device/consumable like sample transfer device, pipette, tube, etc required
    - vi. Name of analyte to be detected
    - vii. Pathogens targeted by the kit

viii. Time taken for testing

ix. Time for result reading and interpretation

x. Manual or automated(equipment)reading

xi. Limit of detection/Limit of Quantification and range of detection

xii. Diagnostic sensitivity

xiii. Diagnostic specificity

xiv. Stability and reproducibility (including data)

xv. Training required for testing (if any)

xvi. If yes, duration

xvii. Details of Cut-off and /or Equivocal Zone for interpretation of test

xviii. Details of cross reactivity, if any

xix. Interpretation of invalid and indeterminate results to be provided

xx. It is recommended to provide data demonstrating accuracy and precision

*****Please mention “Not applicable” against sections not pertaining to the kit.

ANALYTICAL PERFORMANCE EVALUATION OF IN-VITRO DIAGNOSTICS FOR PULMONARY TUBERCULOSIS

LIST of Contributors

A. Working Group:

1. Dr Sivakumar, Scientist E and Head, Department of Bacteriology, ICMR-NIRT, Chennai
2. Dr Madhumathi J, Scientist D, Division of Communicable Diseases (CD), ICMR, New Delhi

B. Review Committee:

1. Dr Amita Jain, Professor and Head, Department of Microbiology, KGMU
2. Dr RM Pandey, ICMR- Dr A.S. Paintal Distinguished Scientist Chair
3. Dr. Camilla Rodrigues, Senior Consultant, P.D. Hinduja Hospital, Mumbai
4. Dr Gita Nataraj, Professor Emeritus, Microbiology, Seth GS Medical College and KEM Hospital, Mumbai
5. Dr Ashutosh Aggarwal, Professor and Head, Pulmonary Medicine, PGIMER
6. Dr Venkataraghava Mohan, Professor & Head, Dept. of Community Health and Development, CMC Vellore
7. Mr. Pramod Meshram, Deputy Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
8. Dr. Sella Senthil, Assistant Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
9. Dr. Md Omair Anwar, Drugs Inspector (Medical Devices), IVD Division, Central Drugs Standard Control Organization, Government of India
10. Dr. Nivedita Gupta, Scientist-G and Head of the Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

Analytical Performance Evaluation of IVD for Pulmonary Tuberculosis

I. Background

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured diagnostic kits appropriate for use in India. This protocol gives the methods to be used for evaluating the analytical performance characteristics of the in-vitro diagnostic test in detecting pulmonary tuberculosis and drug-resistant tuberculosis.

Note: According to CDSCO guidelines, "performance evaluation" refers to "analytical validation" required for obtaining "test license", while “field evaluation” refers to "clinical validation” performed in clinical samples in real world setting.

II. Purpose

To evaluate the performance characteristics of nucleic acid amplification tests (NAAT) forthe diagnosis of pulmonary Mycobacterium tuberculosis (MTB) using irreversibly de-identified leftover archivedor spiked sputum samples.

III. Study Design

Analytical validation of IVD using irreversibly de-identified leftover clinical/spiked samples.

IV. Ethical Considerations

1. Leftover sputum specimens collected for routine diagnostic evaluation from patients who are suspected of having TB shall be used. No additional specimens should be requested.
2. The probability of harm or discomfort anticipated in the research is nil or not expected.
3. Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR’s Guidance on Ethical Requirements for Laboratory Validation Testing, 2024.
4. Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.
5. The protection of privacy of participants should be ensured by using de-identified samples and encrypting the patient identifiers.
6. Respect for the dignity of participants shall be prioritized.

V. Blinding of Laboratory Staff

To ensure the rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff member selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintain the database of results. Staff performing the reference test and the test under evaluation, interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation. The data should be analyzed only by the PI of the evaluating lab. Refer to Fig. 1.

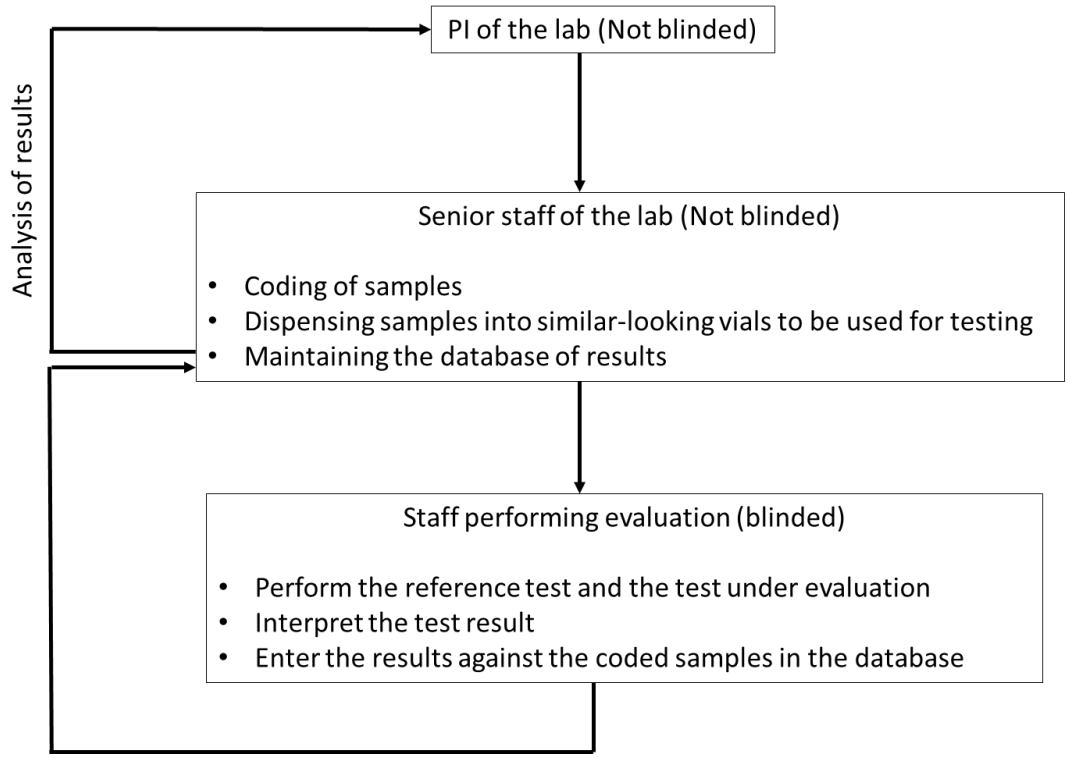


Figure 1 Blinding in evaluation exercise

VI. Procedure

1. Preparation of Evaluation sites/laboratories

C. The laboratory must be approved by the National TB Elimination Program (NTEP).

D. Accreditation for at least one Quality management system [accreditation for Testing Lab / Calibration Lab (ISO/IES 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory].

2. Exclusion

- Extra-pulmonary samples
- Specimens with > 1 freeze-thaw cycle (or according to IFU, if specified)
- Any exclusion criteria stated in the product IFU

3. Reference tests

- ***For detection of MTB:*** Mycobacterium Growth Indicator Tubes (MGIT) liquid culture.
- ***For MDR-TB:*** MGIT drug sensitivity testing (DST)

4. Preparation of samples

- ***For LOD studies - MTBC-negative sputum:*** smear-negative and NAAT-negative sputum should be used for the spiking analytic studies
- ***For analytical sensitivity and specificity:*** Well characterized archived samples (sputum or processed sputum); MTB positives, MTB negatives and Non-Mycobacterium tuberculosis (NTM) samples confirmed by liquid MGIT culture
- ***For drug sensitivity:*** MTB and NTM clinical isolatesthoroughly characterized through MGIT DST and sequencingshould be used.
- For inclusivity/exclusivity, resistance detection, and cross-contamination, mycobacterial strains should be diluted into 7H9 medium at the required concentrations.
- The concentrations (cfu/mL) should be estimated by adjusting the bacterial suspension density to the McFarland standards.

5. Reference Strains

The National Institute for Biological Standards and Control (NIBSC) internal reference standard for Mycobacterium tuberculosis (H37Rv) DNA for Nucleic Acid Amplification Test (NAAT) based assays (NIBSC code: 20/152) will be used for the LOD assay. It was established as the 1st WHO International Standard for Mycobacterium tuberculosis (H37Rv) DNA for NAAT-based assays in 2021. The intended uses of this material are for calibration of secondary or in-house reference

materials used in the assays for the molecular detection of *M. tuberculosis* DNA.It may also be used for assay validation and monitoring the limit of detection of rapid diagnostic tests. This preparation contains an arbitrary unitage of 6.3 log10 (or 2 million) IU per vial.

6. Sample size and sample panel composition

With an anticipated sensitivity of 90% and relative precision of 7%, a minimum of 87 confirmed MTB positive samples by MGIT culture will be required for testing analytical sensitivity. With an anticipated specificity of 95% with 5% relative precision, the minimum sample size required for analytical specificity is 81 confirmed MTB negative samples by MGIT culture. To rule out NTM detection, with an assumed sensitivity of 90% and relative precision of 10%, around 50 confirmed NTM samples may be included to evaluate the index test kit. Hence, approximately 100 confirmed MTB positives, 100 confirmed MTB negatives and 50 NTM samples will be used for pre-validation studies.

The proposed evaluation study will be done using Sputum/MTB isolates stored at the biobank facility of the National TB reference laboratories (NRLs) or the pre-validation labs. The stored sputum/MTB isolate/processed sample/DNA samples will be of the following categories and sub-categories.

Category 1: Positive for MTB by MGIT culture (N = 100) [MTB positives should include equal number of low, medium and high bacterial load samples demonstrated by Xpert MTB]

Category 2: Negative for MTB by MGIT culture (N = 150)

Within the MTB negative group, we propose the following two sub-categories:

i. Negative for all Mycobacteria (N = 100)
ii. Positive for Non-Tuberculous Mycobacterium (N = 50)

Equivalence studies:

If a new specimen or format needs to be evaluated for an already recommended IVD, an equivalence study may be performed with 50 MTB positive and 50 MTB negative specimens to establish relationship between IVD performance in the newly claimed specimen type or new format of same technology. However, if the technology used for the IVD design is different, it shall go through the full process of validation.

Category 3: If resistance detection has to be carried out, within the MTB positive group, we propose to use the following sub-categories:

i. Sensitive to Rifampicin and Isoniazid, individually and combined (N =100) confirmed by Drug susceptibility testing on MGIT liquid culture.
ii. Resistance to both Rifampicin and Isoniazid (N = 100) as detected by Drug susceptibility testing on MGIT liquid culture.
iii. Isoniazid mono-resistance (N =45) as detected by DST on MGIT liquid culture.
iv. Fluroquinolone resistance (N=45) (if applicable for the index test) as confirmed by DST on MGIT liquid culture.

Table 1: Sample size calculation with 95% confidence level

Anticipated Sensitivity	**Relative Precision**	**Sample size**
90%	5%	171
90%	10%	43
90%	7%	87
95%	5%	81
95%	10%	20
95%	7%	41

Analytical sensitivity and specificity:

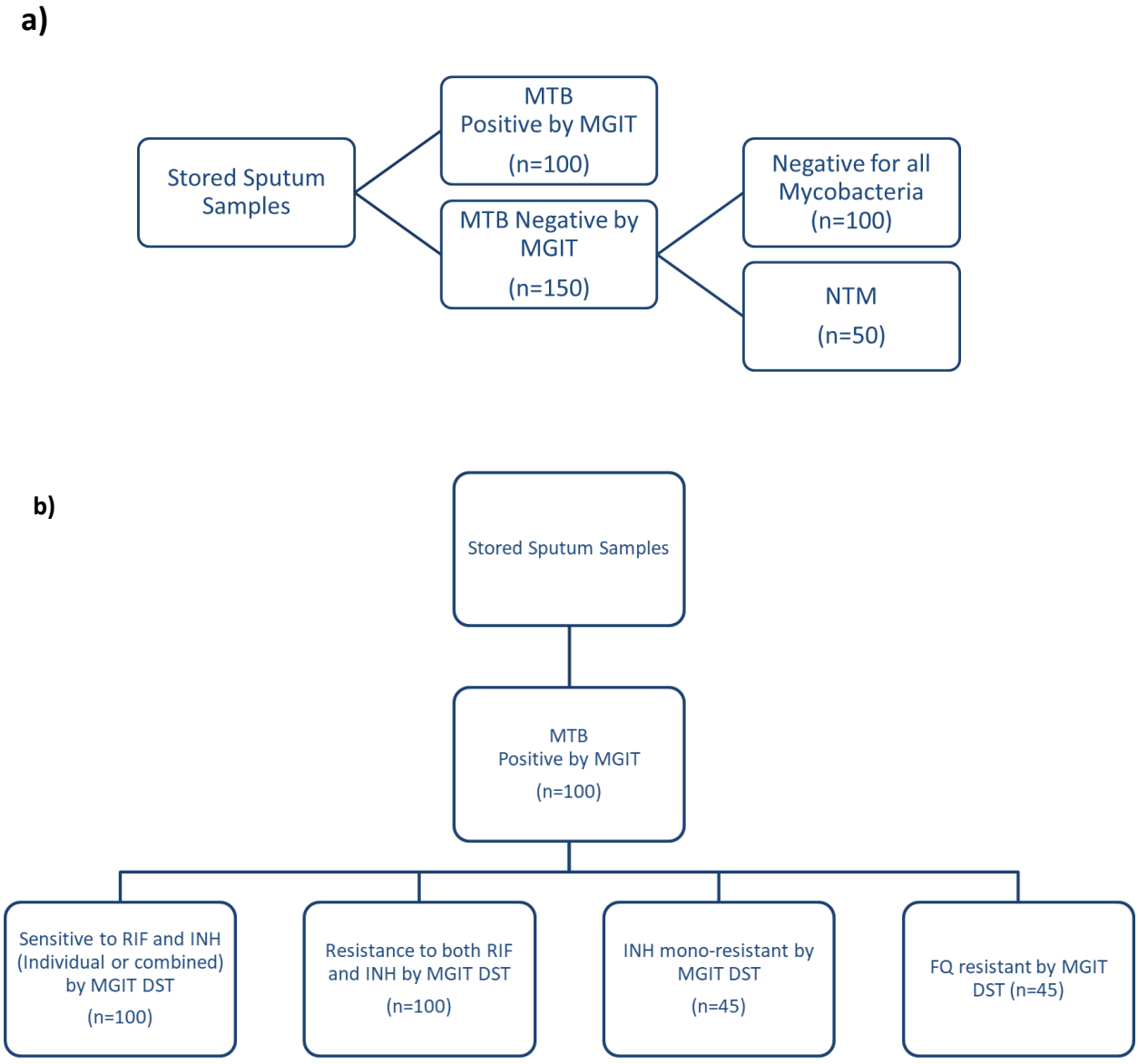


Figure 2. Flowchart for Analytical Performance Evaluation for detection of; a) MTB detection, b) MDR-TB

7. Limit of Detection (LOD) Assay

The 95% LOD is defined as the minimum concentration of bacterium, expressed as CFU/ml or genomic DNA copy numbers/mL, in a sample volume that can be detected in 95% of tests. Finalize the LOD at least one concentration with a hit rate above 95% and two concentrations with hit rates between 10% and 90%. LOD should be always done with NIBSC H37Rv (20/152) standard and only reported in IU/ml or CFU/ml.

Preparation of samples for LOD

1. The volume of sputum required for LOD is based on the IFU (Instruction for use) from the index test manufacturer, which generally varies between 1-2 ml of sputum.
2. A minimum of 200ml of NAAT negative sputum is required for the full LOD studies for a single index test.
3. Sputum samples which are negative by Xpert MTB will be stored at -20^{o}C and once the required amount is obtained the samples will be pooled and tested for MTB using molecular and phenotypic test to prove no growth of MTB in the pooled samples.
4. To perform the assay it may take two weeks to one month based on the multiplicities of test suggested in the IFU after the required volume of sputum is collected.

Spiking of sputum samples

1. The spiked sputum will be used *to determine th*e LOD of the test kit. About 1.8 ml of negative sputum *specimen will* be spiked with 200 ul of the respective diluted suspension series of *M. tuberculosis* H37Rv.
2. These dilutions will be added to the sputum to get the final concentration (10000, 1000, 100, and 10 IU/ml). Before spiking, the culture for CFU wil*l be set up for* the different dilutions.
3. NIBSC reference standard will be reconstituted as directed by NIBSC using 1 mL nuclease free molecular biology grade purified water (MBGPW). From this stock 100 µL will be diluted ½ to get 10,00,000 IU/ml and serially diluted to give 100000, 10000, 1000 and 100 IU/ml with MBGW.
4. Each dilution of the WHO International Standard, will be tested 24 times. The 24 replicates will be performed over at least three days by at least two users and, for low-throughput instruments, on at least three different instruments, or sets of instruments if applicable (e.g., DNA preparation and amplification instruments). For low through-put instruments, the number of testing days may be increased.
5. When *M.tuberculosis* (H37Rv) is used, the 24 replicate tests shall comprise (8 replicate tests on each day for 3 days) of a minimum 8-member 0.5 log 10 dilution panel of a suitable biological reference material (e.g., WHO International Standard) (WHO TSS-17).
6. Each lot shall comprise different production (or manufacturing, purification, etc.) runs of critical reagents. Inter-lot variation must be evaluated by appropriate statistical means.

7. Lowest dilution at which the test detects M.t*b will be deter*mined a LOD, the corresponding CFU will also be counted and reported in terms of CFU/per ml. The LOD will be presented as IU/mL for each dilution.
8. Analytical sensitivity shall be estimated by determining the 95% LOD with 95% confidence intervals (e.g., by probit analysis).
9. If there are more than four invalid results with the same specimen (i.e. dilution) overall, then the specimen will be retested to get at least 20 valid results for each dilution. For tests that include a claim for drug resistance testing, at least 20 valid results (i.e., sensitive or resistant) for each of the claimed drugs should be obtained for each dilution.
10. To arrive at the LOD a probit analysis should be performed, Probit analysis is defined as a specialized form of regression analysis applied to binomial response variables, transforming a concentration-response curve into a straight line for analysis through methods like least squares or maximum likelihood regression. It is primarily used in molecular biology measurement procedures, such as PCR, to determine the detection probability of analytes at various concentrations.

LOD for detection of drug resistance

1. To test the drug resistant MTB strains, well-characterized MTBC strains of known concentration (expressed as CFU/mL) shall be spiked into each claimed MTBC negative specimen type. DR strains shall be characterized by sequencing.
2. Relevant DR strains (as mentioned in table below) shall be spiked into each claimed MTBC-negative specimen type (e.g., raw and/or processed sputum, and each claimed extra-pulmonary specimen).
3. If the assay detects resistance to more than 1 target drug, the LOD for each target drug in addition to a composite resistance LOD, defined as the highest LOD among the tested target, shall be reported.
4. Analytical sensitivity for resistance detection shall be estimated as the lowest number of colony-forming units (CFU) per specimen that can be reproducibly distinguished from negative specimens with 95% confidence.
5. The determination shall comprise 24 replicate tests (8 replicate tests on each of 3 days) of a minimum 8 8-member 0.5log10 dilution panel. The replicate testing shall be conducted on three different days using 2 lots, and at least 2 dilution series shall be tested.

Table 2: Anti-mycobacterial drugs and common mutations

S.No	Drugs	Resistance mutation of strains to be tested
1	Isoniazid	katG_S315T and fabG1_c-15t
2	Rifampicin	rpoB_S450L; rpoB_D435V; rpoB_H445Y; rpoB_H445D; rpoB_D435Y; rpoB_S450W; rpoB_L452P; rpoB_H445L; rpoB_S450F; rpoB_L430P; rpoB_H445R; one rpoC mutation
3	Levofloxacin (CC) LFX2,3	gyrA_A90V, gyrA_D94G, gyrA_D94H, gyrA_D94N, gyrA_D94Y, gyrA_S91P
4	Moxifloxacin (CC and CB)	gyrA_A90V, gyrA_D94G, gyrA_D94H, gyrA_D94N, gyrA_D94Y, gyrA_S91P
5	Bedaquiline	Rv0678_LoF, pepQ_LoF, atpE_p.Ala63Pro
6	Linezolid	rplC_p.Cys154Arg, rrl_n.2814G>T
7	Ethambutol	embB_M306L, embB_M306V, embB_Q497R
8	Delamanid	ddn_LoF, ddn_p.Leu49Pro, fbiC_LoF
9	Pyrazinamide	pncA_ V139A, pncA_ V139G
10	Amikacin	rrs_ A1401G, rrs_ A1401G, rrs_G1484T, eis /promoter_ C-12T, eis /promoter_C-14T
11	Kanamycin	rrs_ A1401G, rrs_ A1401G, rrs_G1484T, eis /promoter_ C-12T, eis /promoter_C-14T
12	Capreomycin	rrs_ A1401G, rrs_ A1401G, rrs_G1484T, eis /promoter_ C-12T, eis /promoter_C-14T
13	Ethionamide	fabG1_c-15t, inhA_S94A, fabG1_ T-8C
14	Pretomanid[#]	ddn_LoF, ddn_p.Leu49Pro, fbiC_LoF
15	Cycloserine	Alr_C-8T, alr_M319T, alr_Y364D, ald_T-32C, ddlA T365A
16	PAS	thyA T22A, folC I43T, folC R49W

8. Reproducibility

Three lots of a test shall be evaluated for lot-to-lot reproducibility. Each lot will comprise different production (or manufacturing, purification, etc.) runs of critical reagents.

- First lot of the assay: should be tested on statistically significant number of positive and negative samples as calculated in the protocol.
- Second lot of the assay: should be tested on 20 samples (10 MGIT positive samples and 10 MGIT negative samples).
- Third lot of the assay: should be tested on 20 samples (10 MGIT positive samples and 25 MGIT negative samples).

Within-run (same operator, same measuring system, same operating conditions, and same location), Between-run, -lot, -day, -site, -operator.

1. Three specimens will be used; MTB sensitive (H37Rv), MTB resistant and MTB

negative.

2. The effect of operator-to-operator variation on IVD performance will be included as part of the precision studies.
3. The nucleic extraction/purification component will also be considered for estimating precision.
4. Contrived specimens will be used (i.e., MTBC strains with specific/most common mutations in the target genes spiked into a clinical matrix claimed in the IFU) for repeatability and reproducibility studies.
5. DR specimens at the concentrations specified for each DRTB (i.e. RR-TB, Hr-TB, MDR-TB, TB resistant to fluoroquinolones) as described in the table on resistance detection.
6. If there are two or more invalid results for the same specimen in the same run, then the run should be repeated for this specimen. Invalid results should be reported.
7. Results will be statistically analyzed by ANOVA or other methods to identify and isolate the sources and extent of any variance.
8. Furthermore, the percentage of correctly identified, incorrectly identified, and invalid results will be compiled for each specimen and separately categorized by site, lot, and other factors.
9. Within-run and within-laboratory reproducibility will be assessed by measuring eight replicates. At least two operators will test a total of 40 positives and 40 negatives with two batches of the kit over a period of 5 days.

9. Inclusivity and exclusivity

1. Inclusivity MTBC stains: For a claim of MTBC detection, the following strains shall be tested: M. tuberculosis, M. bovis, M. bovis BCG, M. africanum, M.microti and M.caprae
2. Exclusivity NTM strains: M. avium, M. kansasii, M. intracellulare
3. Representative MTBC and non-tuberculosis mycobacteria (NTM) strains will be tested in triplicate forinclusivity and exclusivity verification.
4. Resistance detection: For assays with a claim for detection of drug resistance, the applicable specimens from the resistancedetection panelwill be tested in triplicate.
5. The concentration of MTBC isolates used in inclusivity studies will be at levels at or near the specific LOD and will be confirmed by plating/ counting bacterial CFUs (estimated using Truenat).

6. The selection of specific MTBC strains with relevant genetic variations linked to DR will be made to support the claims in the IFU.
7. This will involve testing strains that carry the most common mutations, including associated or interim resistance mutations, covering at least 80% of the resistance mechanisms observed globally for each of the assay target drugs (as shown in table 2).

10. Cross-contamination/carry-over

1. The experiment will allow the determination of the well-to-well or vial-to-vial cross-contamination rate ofhigh-throughput platforms or potential carryover in low-throughput instruments.
2. This will be assessed byalternating one high-positivespecimen with one negative specimen and repeating thissequence twenty times.
3. For high-throughput assays, this will be achieved by alternating high-positive andhigh-negative specimens in the same plate/run.
4. For low-throughput assays, each sequence of highly positivespecimens followed by negative specimens should be done on the same instrument.
5. If more than one instrument isused, each run (i.e same instrument and same day) should include a minimum of 2 sets of alternating high-positive and negative specimens.
6. Contrived specimens prepared by spiking MTBC strains into MTBC negative clinical sputum will be used for these studies.

Note: The strains used for assessment of reproducibility, inclusivity/exclusivity, resistance detection, and carry-over may be commercially acquired or locally prepared, well-characterized strains (by phenotypic DST and sequencing).

11. Resolution of discrepancy:

- The results of MGIT culture should be used to resolve any discrepancy in detection of MTB
- Results of phenotypic DST and sequencing should be used to resolve discrepancy in detection of MDR-TB.

VII. Statistical Analysis Plan

1. The index molecular test should be evaluated for its analytical sensitivity and analytical specificity.

2. 95% Confidence interval should be calculated for each of the parameters.

$$\% \text{ Sensitivity} = \frac{\text{Positives by index test}}{\text{Confirmed positives by MGIT culture}} \times 100 = [a/a+c] * 100$$

$$\% \text{ Specificity} = \frac{\text{Negatives by index test}}{\text{Confirmed negatives by MGIT culture}} \times 100 = [d/b+d] * 100$$

VIII. Acceptance Criteria

Acceptance criteria for Diagnostic tests:

Expected sensitivity: ≥ 90%

Expected specificity: ≥ 95%

Sample Size: ~ 100 confirmed MTB positives (by MGIT culture), ~ 100 confirmed MTB negatives (by MGIT culture) and ~ 50 NTM samples (confirmed by culture and identification)

Acceptance criteria for Screening tests:

Test Type	Minimal Accuracy	Optimal accuracy
High Sensitivity high specificity screening test	90% sensitivity 80% specificity	95% sensitivity 95% specificity
High Sensitivity screening test	90% sensitivity 60% specificity	95% sensitivity 85% specificity
High specificity screening test	60% Sensitivity 98% specificity	70% sensitivity 98% specificity

Source: WHO TPP 2025

IX. Publication Rights

The PI(s) of the evaluating labs shall retain publication rights of the evaluation as lead author(s).

IMPORTANT NOTE

Once a kit is determined to be "Not of Standard Quality", following the procedure outlined in this document, no further requests for repeat testing of that kit will be accepted. Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical

summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

References

1. Technical Specifications Series for Submission to WHO Prequalification – Diagnostic Assessment. TSS 17 In vitro diagnostic medical devices used for the qualitative detection of Mycobacterium tuberculosis complex DNA and mutations associated with drug-resistant tuberculosis.https://iris.who.int/bitstream/handle/10665/366068/9789240055865-eng.pdf.
2. Miotto, P., Tessema, B., Tagliani, E., Chindelevitch, L., Starks, A. M., Emerson, C., Hanna, D., Kim, P. S., Liwski, R., Zignol, M., Gilpin, C., Niemann, S., Denkinger, C. M., Fleming, J., Warren, R. M., Crook, D., Posey, J., Gagneux, S., Hoffner, S., Rodrigues, C., … Rodwell, T. C. (2017). A standardised method for interpreting the association between mutations and phenotypic drug resistance in *Mycobacterium tuberculosis*. *The European respiratory journal*, *50*(6), 1701354. https://doi.org/10.1183/13993003.01354-2017.
3. Indian Catalogue of Mycobacterium tuberculosis Mutations and their Association with Drug Resistance, Version 2.0 2024, ICMR-NIRT. https://www.nirt.res.in/pdf/mutation_catalogue_v2.pdf.
4. Target product profiles for tuberculosis screening tests. Geneva: World Health Organization; 2025. Licence: CC BY-NC-SA 3.0 IGO. Available at: https://iris.who.int/bitstream/handle/10665/382179/9789240113572-eng.pdf?sequence=1.
5. WHO manual for the preparation of secondary reference materials for in vitro diagnostic assays designed for infectious disease nucleic acid or antigen detection: calibration to WHO International Standards. Available at: https://www.who.int/publications/m/item/annex-6-trs-no-1004

PERFORMANCE EVALUATION REPORT FORMAT

Performance Evaluation Report For MTB/MDR-TB Kit

Name of the product (Brand/generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No /Batch No.:		
Product Reference No/ Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number: Issue date: Valid Upto:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak)	
	Negative samples (provide detail: clinical/spiked, including cross reactivity panel)	

Results:

		Reference assay …….…… (MGIT/MGIT DST for RIF/INH/FQ/others)		
		Positive	Negative	Total
Name of MTB or MDR-TB kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

Conclusions:

- Sensitivity, Specificity
- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ..... sample. Results should not be extrapolated to other sample types.)

DISCLAIMERS

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for ……Kit (Lot No……), version ……with the gene targets ……...manufactured by …… (Supplied by …….).

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……

FIELDPERFORMANCE EVALUATION OF IN-VITRO DIAGNOSTICS FOR PULMONARY TUBERCULOSIS

List of Contributors

A. Working Group:

1. Dr Gomathi N. Sivaramakrishnan, Former Scientist, ICMR-NIRT, Chennai
2. Dr Sivakumar, Scientist E and Head, Department of Bacteriology, ICMR-NIRT, Chennai
3. Dr Madhumathi J, Scientist D, Division of Communicable Diseases (CD), ICMR, New Delhi
4. Dr Hansraj Choudhary, Scientist C, Division of Communicable Diseases (CD), ICMR, New Delhi

B. Review Committee:

1. Dr Amita Jain, Professor and Head, Department of Microbiology, KGMU
2. Dr RM Pandey, ICMR- Dr A.S. Paintal Distinguished Scientist Chair
3. Dr. Camilla Rodrigues, Senior Consultant, P.D. Hinduja Hospital, Mumbai
4. Dr Gita Nataraj, Professor Emeritus, Microbiology, Seth GS Medical College and KEM Hospital, Mumbai
5. Dr Ashutosh Aggarwal, Professor and Head, Pulmonary Medicine, PGIMER
6. Dr Venkataraghava Mohan, Professor & Head, Dept. of Community Health and Development, CMC Vellore
7. Mr. Pramod Meshram, Deputy Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
8. Dr. Sella Senthil, Assistant Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
9. Dr. Md Omair Anwar, Drugs Inspector (Medical Devices), IVD Division, Central Drugs Standard Control Organization, Government of India
10. Dr. Nivedita Gupta, Scientist-G and Head of the Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

Field Performance Evaluation of IVD for Pulmonary Tuberculosis

I. Background

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured diagnostic kits appropriate for use in India. This protocol gives the methods to be used for evaluating the clinical performance characteristics of nucleic acid amplification based in-vitro diagnostic test in detecting pulmonary tuberculosis.

Note: According to CDSCO guidelines, "performance evaluation" refers to "analytical validation" required for obtaining "test license", while “field evaluation” refers to "clinical validation” performed in clinical samples in real world setting.

II. Purpose

To evaluate the clinical performance characteristics of nucleic acid amplification tests (NAAT) for diagnosis of pulmonary Mycobacterium Tuberculosis (MTB) using prospectively collected sputum samples in clinical setting.

III. Study Design

Cross-sectional prospective multi-centric diagnostic accuracy study of IVD for detection of pulmonary TB using Mycobacterium Growth Indicator Tube (MGIT) liquid culture as the microbiological reference standard.

IV. Ethical Considerations

1. The study should be compliant to the ICMR’s Guidance on Ethical Requirements for Laboratory Validation Testing, 2024. Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR’s Guidance on Ethical Requirements for Laboratory Validation Testing, 2024. Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.
2. Sputum specimens should be collected, as required for routine diagnostic evaluation, from patients who are suspected of having pulmonary TB as per algorithm. Probability of harm or discomfort anticipated in the research is nil or not expected.
3. Enrolment of subjects should be continued till the sample size is met or till the project duration is completed.
4. If additional sputum sample is obtained, written consent must be obtained as per the ICMR National Ethical Guidelines for Biomedical and Health Research Involving Human Participants. The institutional ethics committee of each participating site should be intimated about the study for necessary approval prior to initiating the study. Assent form

should be collected in addition to Informed Consent in case of adolescents (13 to 16 yrs). For children between 7 and 12 years old, oral assent should be obtained in presence of parent or legal guardian. For children under 7 years old, written informed consent should be obtained from parent or legal guardian.

5. The protection of privacy of research participants will be ensured by encrypting the patient identifiers.
6. Patients shall receive the best possible diagnostic work-up as per the routine practice and the National Tuberculosis Elimination Program (NTEP) guidelines. There should not be delay in sending report due to the study.
7. TB treatment decisions should not be made based on the result of the index test under evaluation, but on the basis of the routine clinical and laboratory methods (smear, solid / liquid culture, standard NAAT results, and clinical work-up).
8. Respect for the dignity of research participants should be prioritized.
9. No compensation shall be provided to the participants since there is no additional cost or travel involved in sample collection for the study. Patients should be compensated for travel and time only if they are asked to pay additional visits exclusively for the sake of the study and not during regular treatment visits.
10. Follow-up visits may be required for a very limited number of discrepant patients to exclude TB.
11. Leftover sputum samples and deposits should be stored for resolving discrepancies. One positive culture and two DNA samples per patient should be stored at -80^{o}C for use later.
12. All the sites should follow up with all study participants till the final diagnosis is made and the patient should be initiated on appropriate treatment as per NTEP norms. Those found to be *M. tuberculosis* complex (MTB)positive by standard NAAT test should be started on anti-tuberculosis treatment (ATT) by medical officer of the study site as per NTEP guidelines.
13. The findings of the study should be made accessible through reports.

V. Blinding of Laboratory Staff

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results.

Staff performing the reference test and the test under evaluation (index test), interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation.

Operators conducting routine laboratory tests (smear, Xpert MTB/RIF, MGIT culture etc) will not participate in the index test evaluation. Instead, dedicated operators, who are not involved in routine testing and are blinded to the routine test results, will perform the index test.The results will be recorded independently for each test without any patient identifiers. The result sheets will be shared with the investigator for result analysis.The data should be analyzed only by the PI of the evaluating lab (Fig. 1).

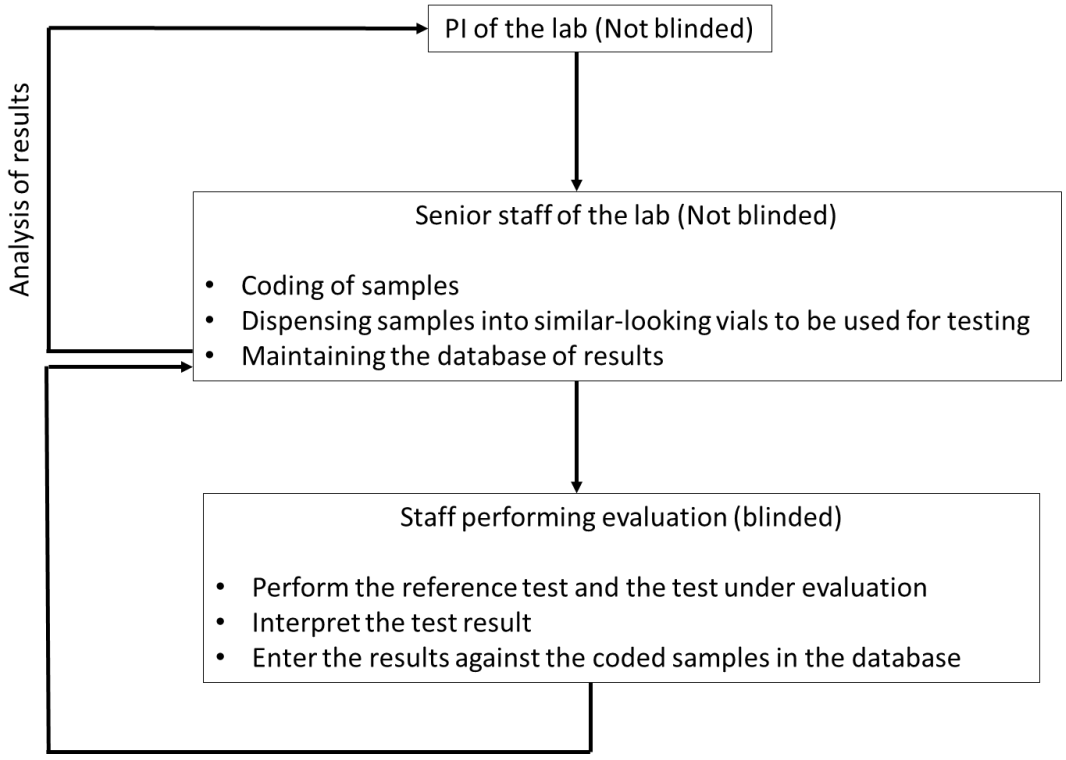


Figure 1 Blinding in evaluation exercise

VI. Procedure

1. Preparation of Evaluation sites/laboratories

- Laboratory must be approved by the National TB Elimination Program (NTEP).
- Accreditation for at least one Quality management system [accreditation for Testing Lab / Calibration Lab (ISO/IES 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory].

- Three or more sites from different geographical regions should perform clinical validation for representation of population in real world setting.

2. Study Participants

Individuals with symptoms of presumptive pulmonary TB attending hospital OPDs/Chest clinics/district microscopy centers (DMCs) and Directly Observed Therapy Short Course (DOTS) centers. All such consecutive cases willing to provide consent will be enrolled in the study.

Definition of Presumptive PTB:

Patients with any of the following symptoms regardless of duration will be considered to have 'presumptive TB': cough for two weeks or more, fever for two weeks or more, night sweats, unintentional weight loss, hemoptysis, chest pain or loss of appetite, with any abnormality in chest radiograph (one or more of the following findings by standardized interpretative criteria: cavitary lesion(s), apical infiltrates, hilar lymphadenopathy, new infiltrates and other suggestive radiological findings).

3. Eligibility of Participants

Inclusion Criteria

1. Individuals positive for TB by smear or any approved NAAT test (Xpert® MTB/RIF)
2. Individuals willing to give consent
3. Individuals who are able and willing to give two good quality mucopurulent sputum samples of ≥ 3 ml

Exclusion criteria

1. Individuals on TB treatment for >96 hrs
2. Individuals not consenting for the study
3. Individuals unable to produce two sputum samples of ≥ 3 ml

4. Reference and Index tests

Reference test: Mycobacterium Growth Indicator Tubes (MGIT) liquid culture

Comparator: NTEP approved NAAT test (Xpert® MTB/RIF)

5. Sample size

The anticipated sensitivity of an index test is 90 % and with absolute 5 % precision, while the anticipated specificity is 99 per cent with 1 % precision. A higher precision for specificity

would be required to minimize false positivity. The minimum sample size requirement has been calculated as ~150 positives and ~470 negatives for MTB by the gold standard culture. With a prevalence of 24 % culture positives among presumptive cases in hospital setting (Penn-Nicholson et al., 2021) and a 5 % loss due to indeterminate results, approximately 610 consecutive cases meeting the inclusion and exclusion criteria would be required to be enrolled for the detection of MTB (Jayaprakasam et al., 2024). Enrolment would be continued till the required number of participants is covered.

The formula for calculating sample size for determining sensitivity/specificity of the index test:

$$N_{Se} = \frac{[Z\ (1-\alpha/2)]^2 * (Se) * (1-Se)]}{d^2}$$

or

$$N_{Sp} = \frac{[Z\ (1-\alpha/2)]^2 * (Sp) * (1-Sp)]}{d^2}$$

N_{Se}: Sample size for estimating sensitivity,

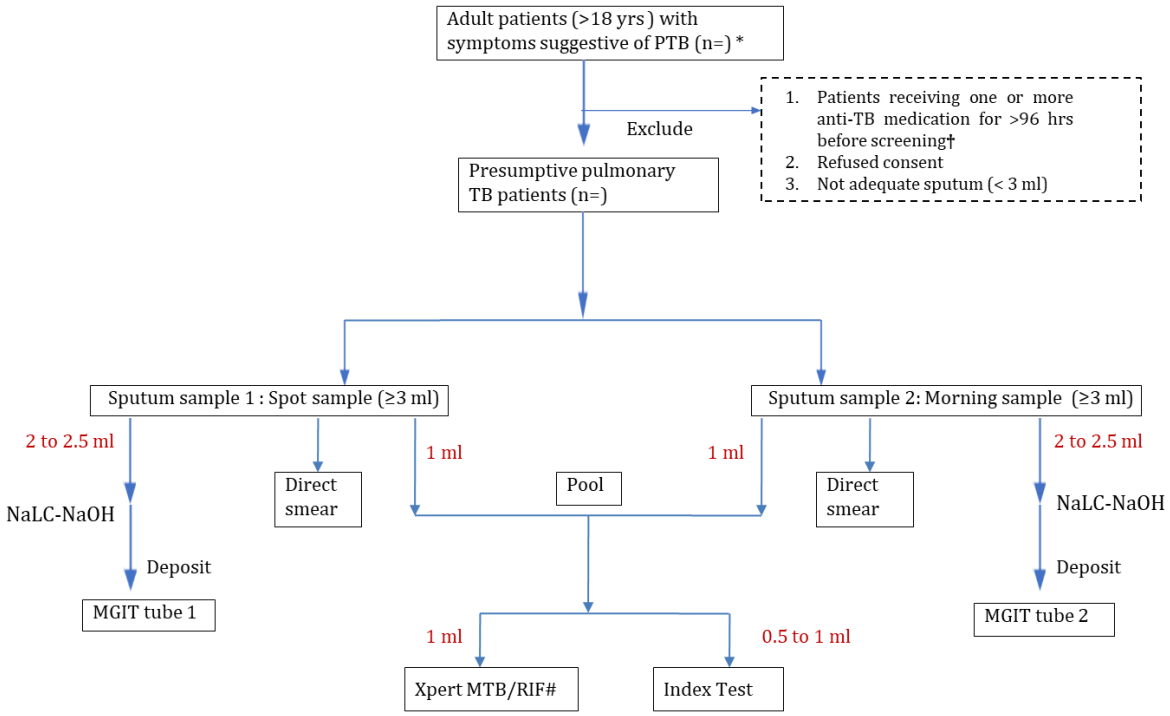
Se: Anticipated sensitivity with reference to culture DST

Sp: Anticipated specificity with reference to culture DST

Z (1-α/2):1.96 for confidence level of 95%

d: Absolute precision

6. Implementation Plan



* ***Screening***: *Medical history & clinical examination as per NTEP guidelines*
† To ensure that dead bacilli are not detected and no treatment failure cases are enrolled
*# **Comparator**: Xpert MTB/RIF*
Storage: *One positive culture and 2 decontaminated samples per patient stored at -80°C for later use. Two DNA samples stored at -20°C for resolution of discrepant results.*

Figure 2 Flowchart for evaluating NAAT test for detection of Mycobacterium Tuberculosis (MTB) among individuals with presumptive pulmonary TB (PTB)

7. Sample collection, processing and storage

1. Two sputum samples each of minimum 3 ml should be collected (one spot and one morning specimen) and sent to laboratory.
2. Approximately 1 ml of sample should be taken from each sample and pooled under sterile conditions (total of 2 ml).
3. Around 1 ml of pooled sample should be tested by the standard NAAT (Xpert MTB/RIF®) and remaining sample used for index test(s).
4. The remaining portion of each sputum sample should be subjected to direct smear and decontamination by NaLC-NaOH method individually.
5. The resultant deposit should be used for inoculation into two MGIT960 tubes.
6. All positive cultures should be identified using rapid Immuno-chromatography test (ICT). (Ideally, positive MGIT tubes are tested within 5 days of instrument positivity. Interpretation of the result should be done within 15 minutes).
7. All sputum samples should be stored at -20^{o}C for later use. Decontaminated sediments and one positive culture per patient should be stored at -80^{o}C, if necessary for later use.
8. Two DNA samples per patient should be stored at -20^{o}C till the end of the study for resolution of discrepant results.
9. The index tests should be carried out as per the algorithm (figure 2) and as per the manufacturers' instructions in the instructions for use (IFU).

All conventional test procedures for smear, culture (solid and liquid) and Xpert MTB will be performed as per NTEP national laboratory guidelines (CTD, 2016; RNTCP 2009) and laboratory manual of ICMR-NIRT (NIRT, 2010). Standard operating procedures for index test(s) will be provided by the manufacturer(s) including use of positive and negative controls. All procedures for preparation of media, reagents, washing, decontamination, disposal and storage will be performed according to the standard operating procedures (SOP) of ICMR-NIRT (NIRT, 2010) and WHO, (WHO, 2022).

8. Laboratory Tests

i. Smear microscopy: Two direct sputum smear
ii. MGIT culture (decontaminated with 1-1.5% final NaOH); Two MGIT tubes (one per specimen) for each patient
iii. Speciation of culture: Rapid immune-chromatographic test (ICT) of MGIT culture

iv. Xpert MTB/RIF (one test per patient)

9. Data Analysis and resolution of discrepancy

i. If the index test produces error or indeterminate results, then only one repeat is allowed. The results of first test and repeat test should be recorded separately.
ii. All Invalids/Indeterminates/errors should be recorded and reported.
iii. A subgroup analysis may be carried out for pediatric population.

10. Quality Control (QC) measures

All sites should ensure high quality of laboratory procedures, data recording and documentation. There should be no deviation from the protocol. All the sites should participate in internal quality control (IQC) and external quality assurance (EQA) for all methods as per the standard manuals of Global Laboratory Initiative (GLI, 2014).

Culture: Positive (Reference strain H37Rv or H37Ra) and negative controls for MGIT and LJ cultures would be tested as per NTEP guidelines. MGIT Time to detection QC for MTB reference strain would be performed every month/new lot of reagents/machine service. Sterility and performance testing of culture media would be performed with every new batch or lot.

Smear: Smear QC should be performed as per NTEP guidelines at regular intervals and with new lot of reagents.

ICT Identification of MTB complex: Culture of *M. tuberculosis* reference strain in MGIT broth should be used as positive control. Culture of Mycobacteria other than tuberculosis (e.g., a well characterized strain of *M. avium* complex/*M.kansasii*) in MGIT broth should be used as negative control. QC for ICT should be performed every 3 months.

Molecular diagnostics: For molecular diagnostics internal quality control includes control supplied by the manufacturer and control prepared by the lab from the previous testing. The internal control should be used whenever batch of test kit changes, machine is serviced, and newly trained person is introduced into the system.

Avoiding Cross-contamination: Unidirectional workflow: The workflow of a molecular lab should be in one direction only. PCR master mix reagents and samples that may contain templates for PCR should be prepared in the pre-PCR room only. Tubes that have undergone amplification in the post-PCR room contain amplicons and will not be opened or introduced in the pre-PCR room. Consumables and PPE (lab coats, gloves, goggles, etc.) that have been used in the post-PCR room should not be placed back in the pre-PCR room without thorough

decontamination. Aerosol resistant pipettes will be used for all procedures and standard aseptic cleaning technique should be carried out before and after PCR for work surface, bench top and equipment.

VII. Statistical Analysis Plan

i. The performance of the diagnostic kits should be evaluated by calculating the sensitivity, specificity, positive predictive value, negative predictive value and accuracy with reference to the gold standard. 95% Confidence interval should be calculated for each of the parameters.
ii. The index molecular test should be evaluated for its performance with reference to the MGIT culture.
iii. Similarly, the performance of standard molecular test (Xpert MTB/RIF) should be estimated with reference to MGIT culture.
iv. The sensitivity and specificity of index test vs MGIT culture should be compared with that of Xpert® MTB/RIF Vs MGIT culture.
v. The agreement between the index test and standard NAAT test (Xpert MTB/RIF) should be calculated with kappa statistic.

VIII. Acceptance Criteria for diagnostic tests

Expected sensitivity: ≥85 ± 2%

Expected specificity: ≥95 ± 2%

Sample size: ~150 MTB positives and ~470 MTB negatives by MGIT culture

For screening tests the acceptability criteria will be as per WHO TPP 2025

Acceptance criteria for Screening tests:

Test Type	Minimal Accuracy	Optimal accuracy
High Sensitivity high specificity screening test	90% sensitivity 80% specificity	95% sensitivity 95% specificity
High Sensitivity screening test	90% sensitivity 60% specificity	95% sensitivity 85% specificity
High specificity screening test	60% Sensitivity 98% specificity	70% sensitivity 98% specificity

Source: WHO TPP 2025

IMPORTANT NOTE

After following due procedure as defined in this document, once any kit is found to be Not of Standard Quality, thereafter, no request for repeat testing of the same kit will be acceptable.

Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

Atleast two different lots or batches should be used for the field validation of any new molecular test.

References

1. Penn-Nicholson, A., Gomathi, S. N., Ugarte-Gil, C., Meaza, A., Lavu, E., Patel, P., Choudhury, B., Rodrigues, C., Chadha, S., Kazi, M., Macé, A., Nabeta, P., Boehme, C., Gangakhedkar, R. R., Sarin, S., Tesfaye, E., Gotuzzo, E., du Cros, P., Tripathy, S., Ruhwald, M., … Members of the Truenat Trial Consortium: (2021). A prospective multicentre diagnostic accuracy study for the Truenat tuberculosis assays. The European respiratory journal, 58(5), 2100526.
2. Jayaprakasam, M., Pandey, R. M., Choudhary, H., Shanmugam, S., Sivaramakrishnan, G. N., & Gupta, N. (2024). Evaluation of molecular diagnostic test for detection of adult pulmonary tuberculosis: A generic protocol. The Indian journal of medical research, 159(2), 246–253.
3. Technical and operational guidelines for tuberculosis control in India 2016. Central TB Division.
4. RNTCP Standard Operating Procedures for Tuberculosis lab for culture and DST, 2009.
5. Standard Operating Procedures (SOP) for Mycobacteriology laboratory, ICMR-NIRT, 2010.
6. Practical manual on tuberculosis laboratory strengthening, 2022 update. Geneva: World Health Organization; 2022. Licence: CC BY-NC-SA 3.0 IGO.
7. Mycobacteriology laboratory manual, Global laboratory initiative, First edition, April 2014, Stop TB Partnership.

PERFORMANCE EVALUATION REPORT FORMAT

Performance Evaluation Report ForMTB Kit

Name of the product (Brand/generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No /Batch No.:		
Product Reference No/Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number: Issue date: Valid Upto:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak)	
	Negative samples (provide detail: clinical/spiked, including cross reactivity panel)	

Results:

Test	Number of samples tested	Positive	Negative	Invalids/ Indeterminates/Error/ Contamination (culture)
Smear				
MGIT culture				
Xpert MTB/RIF				
New MTB kit				

		Reference assay ……

		(MGIT culture)		
		Positive	Negative	Total
Name of MTB kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

Conclusions:

- Sensitivity, specificity
- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ..... sample. Results should not be extrapolated to other sample types.)

DISCLAIMERS

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for ……Kit (Lot No……), version ……with the gene targets ……...manufactured by …… (Supplied by …….).

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……

FIELD PERFORMANCE EVALUATION OF IN-VITRO DIAGNOSTICS FOR PULMONARY DRUG RESISTANT TUBERCULOSIS

List of Contributors

A. Working Group:

1. Dr Sivakumar, Scientist E and Head, Department of Bacteriology, ICMR-NIRT, Chennai
2. Dr Joy Sarojini Michael, Professor, CMC, Vellore
3. Dr Shubhada Shenai, Senior Scientist, FIND
4. Dr Madhumathi J, Scientist D, Division of Communicable Diseases (CD), ICMR, New Delhi

B. Review Committee:

1. Dr Amita Jain, Professor and Head, Department of Microbiology, KGMU
2. Dr RM Pandey, ICMR- Dr A.S. Paintal Distinguished Scientist Chair
3. Dr. Camilla Rodrigues, Senior Consultant, P.D. Hinduja Hospital, Mumbai
4. Dr Gita Nataraj, Professor Emeritus, Microbiology, Seth GS Medical College and KEM Hospital, Mumbai
5. Dr Ashutosh Aggarwal, Professor and Head, Pulmonary Medicine, PGIMER
6. Dr Venkataraghava Mohan, Professor & Head, Dept. of Community Health and Development, CMC Vellore
7. Mr PramodMeshram, Deputy Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
8. D. Sella Senthil, Assistant Drugs Controller, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India
9. Dr. Md Omair Anwar, Drugs Inspector (Medical Devices), IVD Division, Central Drugs Standard Control Organization, Government of India
10. Dr Nivedita Gupta, Scientist-G and Head of the Division of Communicable Diseases, ICMR Headquarters, Department of Health Research, Ministry of Health and Family Welfare, Government of India

Field Performance Evaluation of IVD for Pulmonary DR-TB

I. Background

CDSCO and ICMR, New Delhi, have aimed at facilitating the availability of Quality-Assured diagnostic kits appropriate for use in India. This protocol gives the methods to be used for evaluating the clinical performance characteristics of the in-vitro diagnostic test in detecting pulmonary drug resistant tuberculosis (DR-TB).

Note: According to CDSCO guidelines, "performance evaluation" refers to "analytical validation" required for obtaining "test license", while “field evaluation” refers to "clinical validation” performed in clinical samples in real world setting.

II. Purpose

To evaluate the clinical performance characteristics of nucleic acid amplification tests (NAAT) fordiagnosis of pulmonary drug resistant tuberculosis (DR-TB) using prospectively collected sputum samples in clinical settings.

Primary Objectives

1. To determine the diagnostic accuracy of new multi-drug resistant (MDR) NAAT test against culture based drug sensitivity testing (DST) in detecting first line drug resistance [Rifampicin (RIF), Isoniazid (INH)] among the microbiologically confirmed TB patients (positive by smear or NAAT test).
2. To determine the diagnostic accuracy of new NAAT test against culture-based drug sensitivity testing (DST) in detecting fluroquinolone drug resistance (FQ) among the microbiologically confirmed TB patients (positive by smear or NAAT test).

III. Study Design

Cross-sectional prospective multi-centric diagnostic accuracy study of IVD for detection of pulmonary drug resistant TB,using Mycobacterium Growth Indicator Tube culture and drug sensitivity testing (MGIT-DST) as the microbiological reference standard.

IV. Ethical Considerations

1. The study should be compliant to the ICMR’s Guidance on Ethical Requirements for Laboratory Validation Testing, 2024. Performance evaluation activities using irreversibly de-identified leftover clinical samples are exempt from ethics approval as per ICMR’s Guidance on Ethical Requirements for Laboratory Validation Testing,

2024. Investigators are required to submit a self-declaration form, as outlined in the ICMR guidelines, to the institutional authorities and ethics committee for information.

2. Sputum specimens should be collected, as required for routine diagnostic evaluation, from patients who are suspected of having pulmonary TB as per algorithm. Probability of harm or discomfort anticipated in the research is nil or not expected.
3. Enrolment of subjects should be continued till the sample size is met or till the project duration is completed.
4. If additional sputum sample is obtained, written consent must be obtained as per the ICMR National Ethical Guidelines for Biomedical and Health Research Involving Human Participants. The institutional ethics committee of each participating site should be intimated about the study for necessary approval prior to initiating the study. Assent form should be collected in addition to informed consent in case of adolescents (13 to 16 yrs). For children between 7 and 12 years old, oral assent should be obtained in presence of parent or legal guardian. For children under 7 years old, written informed consent should be obtained from parent or legal guardian.
5. The protection of privacy of research participants will be ensured by encrypting the patient identifiers.
6. Patients shall receive the best possible diagnostic work-up as per the routine practice and the National Tuberculosis Elimination Program (NTEP) guidelines. There should not be delay in sending report due to the study.
7. TB treatment decisions should not be made based on the result of the index test under evaluation, but on the basis of the routine clinical and laboratory methods (smear, solid / liquid culture, standard NAAT results, and clinical work-up).
8. Respect for the dignity of research participants should be prioritized.
9. No compensation shall be provided to the participants since there is no additional cost or travel involved in sample collection for the study. Patients should be compensated for travel and time only if they are asked to pay additional visits exclusively for the sake of the study and not during regular treatment visits.
10. Follow-up visits may be required for a very limited number of discrepant patients to exclude TB.
11. Leftover sputum samples and deposits should be stored for resolving discrepancies. One positive culture and two DNA samples per patient should be stored at -80^{o}C for use later.
12. All the sites should follow up with all study participants till the final diagnosis is made

and the patient should be initiated on appropriate treatment as per NTEP norms. Those found to be *M. tuberculosis* complex (MTB)positive by standard NAAT test should be started on anti-tuberculosis treatment (ATT) by medical officer of the study site as per NTEP guidelines.

13. The findings of the study should be made accessible through reports.

V. Blinding of Laboratory Staff

To ensure rigor of the evaluation process, laboratory staff performing the evaluation should be blinded to the status of the clinical samples. The PI of the evaluation exercise should remain unblinded, i.e., privy to the status of the samples. Another senior laboratory staff selected by the PI may remain unblinded and carry out coding of samples and dispensing them into similar-looking vials to be used for testing, and maintaining the database of results.

Staff performing the reference test and the test under evaluation (index test), interpretation of the test result, and entering the results against the coded samples in the database, should remain blinded to the status of samples till the completion of evaluation.

Operators conducting routine laboratory tests (GeneXpert MTB/RIF, MGIT DST, LPA etc.) will not participate in the index test evaluation. Instead, dedicated operators, who are not involved in routine testing and are blinded to the routine test results, will perform the index test. The results will be recorded independently for each test without any patient identifiers. The result sheets will be shared with the investigator for result analysis.The evaluation study data should be analyzed only by the PI of the evaluating lab (Fig. 1).

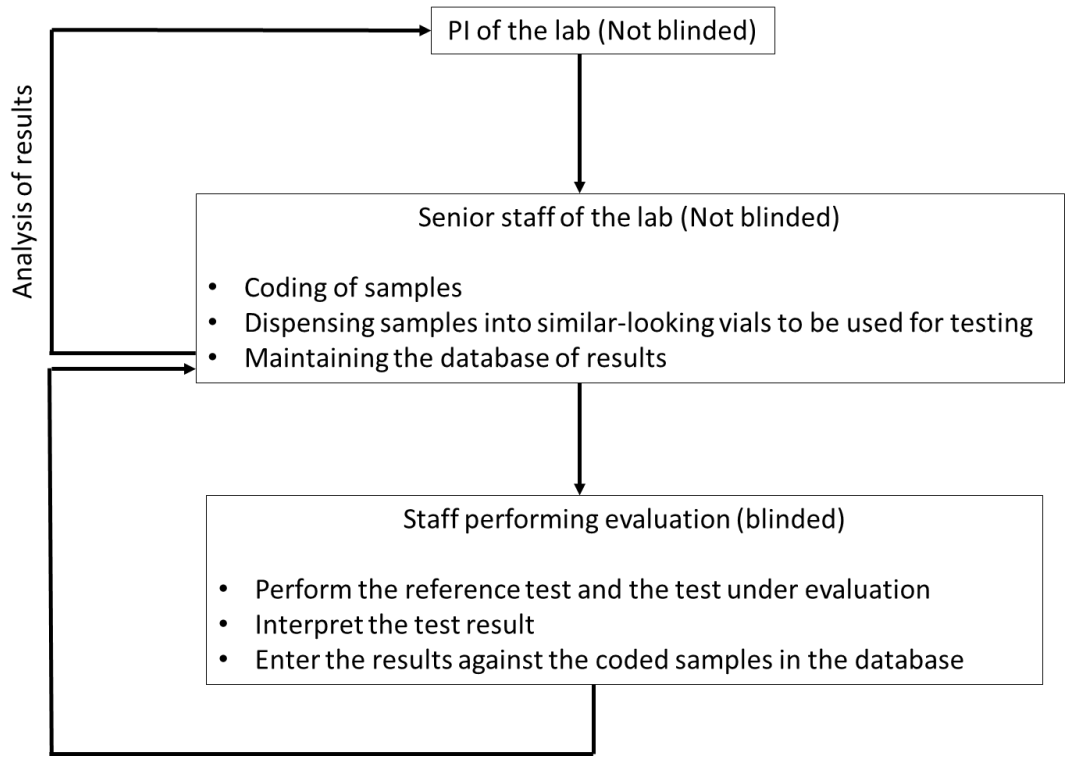


Figure 1 Blinding in evaluation exercise

VI. Procedure

1. Preparation of Evaluation sites/laboratories

- Laboratory must be approved by the NTEP.
- Accreditation for at least one Quality management system [accreditation for Testing Lab / Calibration Lab (ISO/IES 17025), Medical Lab (ISO 15189), PT provider ISO/IEC 17043 or CDSCO approved Reference laboratory].
- Three or more sites from different geographical regions should perform clinical validation for representation of population in real world setting.

2. Study Participants

People with microbiologically confirmed pulmonary TB by smear and/or NTEP approved NAAT test attending hospital OPDs/Chest clinics/district microscopy centers (DMCs) and Directly Observed Therapy Short Course (DOTS) centers. All such consecutive cases (not currently receiving ATT) and willing to provide consent should be enrolled in the study.

3. Eligibility of Participants

Inclusion criteria for testing First Line Drugs

i. Individuals positive for TB by smear or any approved NAAT test (Xpert®

MTB/RIF) and not receiving ATT

ii. Individuals willing to give consent

iii. Individuals who are able and willing to give two good quality mucopurulent sputum samples of ≥ 3 ml

Exclusion criteria

i. Individuals on TB treatment for >10 days

ii. Individuals not consenting for the study

iii. Individuals unable to produce two sputum samples of ≥ 3 ml

4. Reference and Index tests

	Index test	**Reference Test**	**Comparator**
First Line Drug Resistance	New NAAT test for RIF/INH	MGIT Culture DST for RIF and INH	FL-LPA: GenoType MTBDRplus
Second Line Drug Resistance	New NAAT test for FQ	MGIT Culture DST for Moxifloxacin (0.25, 1 mg) and Levofloxacin (1 mg)	SL-LPA: GenoType MTBDRsl

5. Sample size

Sample size for RIF and INH resistance among TB patients

The expected sensitivity of the index test is about 90% with 5 % precision and the expected specificity is 95% with 5% precision. With a confidence interval of 95 % and assuming 10 % loss due to indeterminate results, the sample size required is estimated to be approximately **200**patient's positive each for INH and RIF resistance either alone or in combination. The average prevalence of Isoniazid and Rifampicin are ~18 % and 7.3 % respectively, among the new and previously treated TB patients combined together(Report of drug resistance survey, 2014-16). The number needed to screen to obtain 200 drug resistant cases will be approximately 1111 for INH resistance and 2857 for RIF resistance. The participants will be enrolled till the required sample size is achieved for INH and RIF resistance.

The expected sensitivity of the index test for detecting FQ resistance is 85 % with 7 % precision and the expected specificity is 95 % with 5 % precision. Assuming 10 % loss, the sample size required is 111 FQ resistant cases. The prevalence of FQ resistance among TB patients is ~3

% (Report of drug resistance survey, 2014-16). Hence the number needed to screen will be approximately 3333. The participants will be enrolled till the required sample size is achieved for FQ resistance. Table 1 shows sample sizes required for RIF, INH and FQ drug resistance.

Table 1. Sample sizes for RIF, INH and FQ Drug Resistance

	Assumptions for Sensitivity	**Assumptions for Specificity**
Sensitivity/Specificity of the new test (%)	90	95
Relative precision (d) (%)	5	5
Desired confidence level (1- alpha) %	95	95
Number of drug resistance (INH and RIF) cases required	178	84
Number of drug resistant cases required with 10 % loss due to indeterminate results	~200	~93
Number needed to be screened assuming a combined weighted average prevalence of ~18 % for INH resistance among the new and previously treated TB patients	1111	517
Number needed to be screened assuming a combined weighted average prevalence of ~7 % for RIF resistance among the new and previously treated TB patients	2857	1329

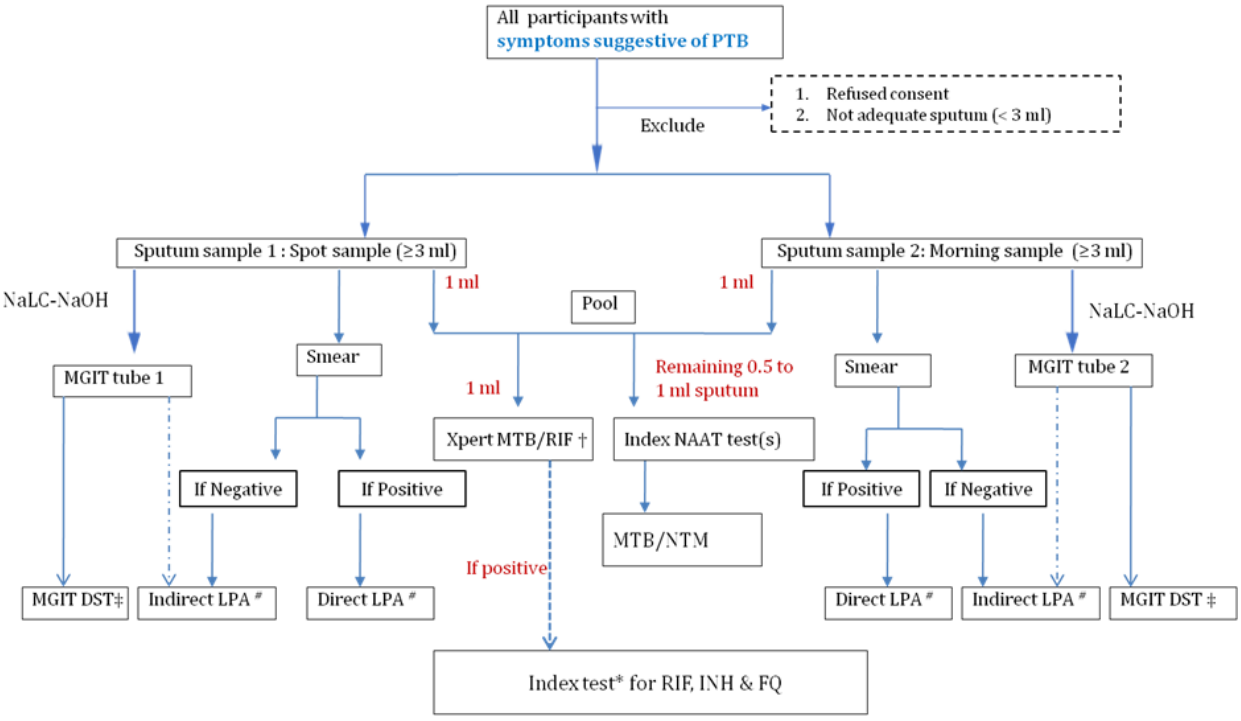
Other disease controls (to check cross-reactivity in real patients)

Include people with common alternative diagnoses to mirror programmatic reality and probe false positives. This subset helps characterize clinical exclusivity beyond simple "TB-negative" status:

i. Non-Tuberculous Mycobacteria (Culture or PCR confirmed): ~30
ii. Other respiratory diseases [e.g., bacterial pneumonia, chronic obstructive pulmonary disease (COPD), lung cancer, chronic fungal (like Histoplasmosis or Aspergillosis)]: ~30 patients combined.

6. Implementation Plan

The samples will be collected and tested as per the routine practice for smear, Xpert MTB/RIF®, LPA, MGIT culture and DST. The samples with positive result for MTB either in smear or NAAT test should be tested for first line and second line drug resistance (RIF, INH and FQ).



** **Index test RIF and INH:** Samples tested positive by either smear or Xpert will be tested by Index test for drug resistance*
*$ **Index Test**: New NAAT Test under evaluation*
*# **LPA**: Any one positive sample will be used for LPA- Direct LPA if smear positive and indirect LPA if smear negative and culture positive.*
*‡ **MGIT DST**: Any one positive culture (tube 1 or 2) will be used for DST*
***Storage**: Leftover sputum samples and DNA elutes to be stored at -20°C, One positive culture and two decontaminated sediments per patient will be stored at -80°C for later use*

Figure 2. Flowchart for evaluating IVDs for testing drug resistance to RIF, INH and FQ among pulmonary TB (PTB) patients

7. Sample collection, processing and storage

1. Two sputum samples each of minimum 3 ml should be collected (one spot and one morning specimen) and sent to laboratory.
2. Approximately 1 ml of sample should be taken from each sample and pooled under sterile conditions (total of 2 ml).
3. Around 1 ml of pooled sample should be tested by the standard NAAT (Xpert MTB/RIF®) and remaining sample used for index test(s).
4. The remaining portion of each sputum sample should be subjected to direct smear and

decontamination by NaLC-NaOH method individually.

5. All smear positive or NAAT positive samples will be tested by Line Probe Assay (LPA).
6. The resultant deposit should be used for inoculation into two MGIT960 tubes.
7. All positive cultures should be identified using rapid Immuno-chromatography test (ICT). (Ideally, positive MGIT tubes are tested within 5 days of instrument positivity. Interpretation of the result should be done within 15 minutes).
8. The positive cultures should be tested for drug sensitivity.
9. All sputum samples should be stored at -20^{o}C for later use. Decontaminated sediments and one positive culture per patient should be stored at -80^{o}C, if necessary for later use.
10. Two DNA samples (one DNA sample extracted for index test and one for LPA) per patient should be stored at -20^{o}C till the end of the study for resolution of discrepant results.
11. The index tests should be carried out as per the algorithm (figure 2) and as per the manufacturers' instructions in the instructions for use (IFU).

All conventional test procedures for smear, culture (solid and liquid) and Xpert MTB will be performed as per NTEP national laboratory guidelines (CTD, 2016; RNTCP 2009) and laboratory manual of ICMR-NIRT (NIRT, 2010). Standard operating procedures for index test(s) will be provided by the manufacturer(s) including use of positive and negative controls. All procedures for preparation of media, reagents, washing, decontamination, disposal and storage will be performed according to the standard operating procedures (SOP) of ICMR-NIRT (NIRT, 2010) and WHO, (WHO, 2022).

8. Laboratory Tests

v. Smear microscopy: Two direct sputum smear
vi. MGIT culture (decontaminated with 1-1.5% final NaOH); Two MGIT tubes (one per specimen) for each patient
vii. MGIT drug sensitivity testing (DST) for Rif, INH: Drug sensitivity testing will be carried out from any one positive MGIT culture.
viii. MGIT drug sensitivity testing for moxifloxicin (0.25 mg and 1 mg) and levofloxacin (1 mg). Drug sensitivity testing should be carried out in from any one positive MGIT culture.
ix. Speciation of culture: Rapid immunochromatographic test (ICT) of MGIT culture
x. LPA: LPA shall be carried out as per routine practice and as per NTEP guidelines. Direct LPA should be carried out from any one smear positive sample. If the sample is smear

negative and culture positive, indirect LPA should be carried out from culture. First line LPA (FL-LPA) will be carried out (Rif and INH resistance)

xi. Xpert MTB/RIF (one test per patient)

9. Index test

i. Index test will be performed as per manufacturer's instructions following blinded study protocols.

ii. At least 2 different lots of reagents should be tested across the study population to demonstrate consistency of test performance and minimize lot-related bias.

iii. The results of the index test will not be disclosed to study participants or clinicians and will not be used to guide treatment decisions.

10 . Data Analysis and resolution of discrepancy

iv. If the index test produces error or indeterminate results, then only one repeat is allowed. The results of first test and repeat test should be recorded separately. All Invalids/Indeterminates/errors should be recorded and reported.

v. Results for new patients and previously treated patients should be entered separately. Result analysis will be carried out for these two populations separately as well as combined.

vi. A subgroup analysis may be carried out for pediatric population.

11. Quality Control (QC) measures

All sites should ensure high quality laboratory procedures, data recording and documentation. There should be no deviation from the protocol. All the sites should participate in internal quality control (IQC) and external quality assurance (EQA) for all methods as per the standard manuals of Global Laboratory Initiative (GLI, 2014).

Culture: Positive (Reference strain H37Rv or H37Ra) and negative controls for MGIT and LJ cultures would be tested as per NTEP guidelines. MGIT Time to detection QC for MTB reference strain would be performed every month/new lot of reagents/machine service. Sterility and performance testing of culture media would be performed with every new batch or lot.

Drug sensitivity testing (DST): Standard ATCC strains should be used for each drug as reference control. QC should be performed whenever a new batch of drugs is prepared, after servicing of the instrument and after long gap of setting up DST.

Molecular diagnostics: For molecular diagnostics internal quality control includes control

supplied by the manufacturer and control prepared by the lab from the previous testing. The internal control should be used whenever batch of test kit changes, machine is serviced, and newly trained person is introduced into the system.

VII. Statistical Analysis Plan

i. The performance of the diagnostic kits should be evaluated by calculating the sensitivity, specificity, positive predictive value, negative predictive value and accuracy with reference to the gold standard. 95% Confidence interval should be calculated for each of the parameters.
ii. The index molecular test will be evaluated for its performance with reference to MGIT DST (for RIF/INH/FQ).
iii. Similarly, the performance of NTEP approved molecular test (Xpert MTB/RIF and LPA) should be estimated with reference to MGIT DST.
iv. The agreement between the index test and molecular test for drug resistance (LPA) should be calculated using kappa statistic.

VIII. Acceptance Criteria

Expected minimal sensitivity for MTB and Drug Resistant TB: ≥85 ± 2%
Expected minimal specificityfor MTB and Drug Resistant TB: ≥95 ± 2%
Sample size: ~200 positives for each drug resistance (RIF or INH or FQ etc) (either alone or in combination) and ~ 100 negatives for each drug resistance (RIF or INH or FQ etc).

IMPORTANT NOTE

Once a kit is determined to be "Not of Standard Quality", following the procedure outlined in this document, no further requests for repeat testing of that kit will be accepted. Any request of re-validation from the same manufacturer for the same test type will only be entertained after a minimum of 3 months and only if a high-level technical summary of modifications or functional improvements to the kit design is submitted, without explicit disclosure of proprietary information.

Clinical samples are precious, therefore, repeat evaluation of a kit using the same/ different well-characterized sample panel at a different laboratory may be considered only for kits which claim high performance characteristics (sensitivity and specificity 95% and above), but which fail the performance evaluation by a margin of 5%.

Atleast two different lots or batches should be used for the field validation of any new molecular test.

References

1) Report of the first national anti-tuberculosis drug resistance survey India, 2014-2016.
2) Technical and operational guidelines for tuberculosis control in India 2016. Central TB Division.
3) RNTCP Standard Operating Procedures for Tuberculosis lab for culture and DST, 2009.
4) Standard Operating Procedures (SOP) for Mycobacteriology laboratory, ICMR-NIRT, 2010.
5) Practical manual on tuberculosis laboratory strengthening, 2022 update. Geneva: World Health Organization; 2022. Licence: CC BY-NC-SA 3.0 IGO.
6) Mycobacteriology laboratory manual, Global laboratory initiative, First edition, April 2014, Stop TB Partnership.

PERFORMANCE EVALUATION REPORT FORMAT

Performance Evaluation Report For MDR-TB Kit

Name of the product (Brand/generic)		
Name and address of the legal manufacturer		
Name and address of the actual manufacturing site		
Name and address of the Importer		
Name of supplier: Manufacturer/Importer/Port office of CDSCO/State licensing Authority		
Lot No /Batch No.:		
Product Reference No/Catalogue No		
Type of Assay		
Kit components		
Manufacturing Date		
Expiry Date		
Pack size (Number of tests per kit)		
Intended Use		
Number of Tests Received		
Regulatory Approval: Import license / Manufacturing license/ Test license License Number: Issue date: Valid Upto:		
Application No.		
Sample Panel	Sample type	
	Positive samples (provide details: strong, moderate, weak)	
	Negative samples (provide detail: clinical/spiked, including cross reactivity panel)	

Results:

Test	**Number of samples tested**	**Positive**	**Negative**	**Invalids/Indeterminates/ Error/Contamination (culture)**
Smear				
MGIT culture				
Xpert MTB/RIF				
	Number of samples tested	**Sensitive**	**Resistant**	
FL LPA – RIF				
FL LPA - INH				
SL LPA- FQ				
MGIT-DST- RIF				
MGIT-DST-INH				

MGIT-DST-FQ				
New IVD- RIF				
New IVD-INH				
New IVD-FQ				

		Reference assay ……..…… (MGITDST – RIF/INH/FQ)*		
		Positive	Negative	Total
Name of MDR-TB kit	Positive			
	Negative			
	Total			

	Estimate (%)	95% CI
Sensitivity		
Specificity		

***Report RIF/INH/FQ as separate tables**

Conclusions:

- Sensitivity, specificity
- Performance: **Satisfactory / Not satisfactory**

(Sensitivity and specificity have been assessed in controlled lab setting using kits provided by the manufacturer from the batch mentioned above using ….. sample. Results should not be extrapolated to other sample types.)

DISCLAIMERS

1. This validation process does not approve / disapprove the kit design
2. This validation process does not certify user friendliness of the kit / assay

Note: This report is exclusively for ……Kit (Lot Nos.……), version ……with the gene targets ……...manufactured by …… (Supplied by …….).

Evaluation Done on ……

Evaluation Done by …….

Signature of Director/ Director-In-charge …… Seal ……

**********End of the Document**********


INDIAN COUNCIL OF MEDICAL RESEARCH
NEW DELHI


icmr
INDIAN COUNCIL OF
MEDICAL RESEARCH
Serving the nation since 1911


CENTRAL DRUGS STANDARD CONTROL ORGANIZATION
CDSCO
CDSCO
MINISTRY OF HEALTH • GOVERNMENT OF INDIA