Asian Harmonization Working Party
WORKING TOWARDS MEDICAL DEVICE HARMONIZATION IN ASIA

FINAL DOCUMENT

Title: Clinical Evidence for IVD Medical Devices - Clinical Performance Studies for In Vitro Diagnostic Medical Devices

Authoring Group: Work Group 2 - Pre-market: IVDD
Work Group 5 - Clinical Evidence for Performance and Safety

Date: 21 May 2021

Dr. Wen-Wei TSAI
Chair, Working Group 2

Mr. Bin Imran FIKRIANSYAH
Chair, Working Group 5

Acknowledgements

This Guidance document was prepared by a sub-group of Working Groups 2 and 5 of AHWP. We wish to acknowledge the contributions of project team members: Ms. Shelley Tang, Ms. Judy Cheung, Dr. Tzu Ling Tseng, Dr. Veronica Wong, Mr. Christopher Chan of Work Group 2, and Ms. Sumati Randeo, Ms. Mie Ohama of Work Group 5.

Table of Contents

Preface

This document is produced by the Asian Harmonization Working Party, based on Global Harmonization Task Force Final Document GHTF/SG5/N8: 2012 of GHTF Study Group 5. The document is intended to provide non-binding guidance for use in the regulatory system of in vitro diagnostic (IVD) medical devices, and has been subject to consultation throughout its development.

1.0 Introduction

When placing an IVD medical device on the market the manufacturer must have demonstrated through the use of appropriate conformity assessment procedures that the device complies with AHWP's Essential Principles of Safety and Performance of Medical Devices (EPs). As IVD medical devices are used for the examination of specimens derived from the human body, the characteristics of clinical evidence are different from medical devices other than IVD medical devices.

Generally, from a clinical perspective, it is expected that the manufacturer has demonstrated the IVD medical device achieves its intended performance during normal conditions of use by the intended user in the intended environment (e.g. laboratories, physician's offices, healthcare centers, home environments) and in the intended use population.

The objective of a clinical performance study is to evaluate whether the IVD medical device is suitable (i.e. meets the relevant Essential Principles of Safety and Performance) for the purpose(s) and the population(s) for which it is intended, when this cannot be addressed with the analytical performance data, literature and/or experience gained by routine diagnostic testing.

In general, clinical performance studies must be designed according to, and take into account scientific principles underlying the collection of clinical performance data along with accepted operational and ethical standards surrounding the use of human subjects. The clinical performance study objectives and design should be documented in a clinical performance study protocol. The data collection process must ensure patient safety and data integrity along the entire process of the study.

2.0 Scope

The primary purpose of this document is to provide guidance for manufacturers of IVD medical devices, Regulatory Authorities (RA) and Conformity Assessment Bodies (CAB) in relation to:

- the selection of a clinical performance study design;
- considerations to be made when undertaking a clinical performance study; and
- protocol, conduct and report of a clinical performance study.

Given the wide diversity of IVD medical devices and their associated risks, this document is not intended to provide comprehensive guidance for clinical performance studies of specific IVD medical devices.

Clinical performance studies are typically performed in the pre-market phase. However, this document may also apply to studies conducted in the post-market phase. Analytical performance studies are beyond the scope of this document.

NOTE: This document should not be used in isolation but read together with the documents *'Clinical evidence for IVD medical devices – Key Definitions and Concepts' and 'Clinical evidence for IVD medical devices – Scientific Validity Determination and Performance Evaluation'* .

3.0 References

AHWP/WG2-WG1/F001:2016 *Definition of the Terms "Medical Device" and "In Vitro Diagnostic (IVD) Medical Device"*

AHWP/WG1a/F002:2013 *Essential Principles of Safety and Performance of IVD Medical Devices*

AHWP/WG2/F001:2016 *Principles of In Vitro Diagnostic Medical Devices Classification.*

AHWP/WG2/F002:2016 *Principles of Conformity Assessment for in vitro Diagnostic (IVD) Medical Devices*

AHWP/WG2/F003:2016 *Submission Dossier for Demonstrating Conformity to the Essential Principles of Safety and Performance of In Vitro Diagnostic Medical Devices*

AHWP/WG5/F003:2015 *Clinical Evidence for IVD Medical Devices – Key Definitions and Concepts*

AHWP/WG5/F004:2015 *Clinical Evidence for IVD Medical Devices – Scientific Validity Determination and Performance Evaluation*

ISO 20916:2019 *In vitro diagnostic medical devices – Clinical performance studies using specimens from human subjects – Good study practice*

World Medical Association – *Declaration of Helsinki – Ethical principles for medical research involving human subjects*

Guidance for the design and reporting of studies evaluating the clinical performance of tests for present or past SARS-CoV-2 infection.; BMJ 2021;372:n568

Clinical Evidence Requirements for CE certification under the Diagnostic Regulation in the European Union ; MedTech Europe

Guidance for Industry and FDA Staff - Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests ; US Food and Drug Administration (FDA)

4.0 Definitions

Clinical Performance of an IVD Medical Device

Ability of an IVD medical device to yield results that are correlated with a particular clinical condition or physiological/pathological process/state in accordance with the intended use (clinical test purpose, target population and intended user).

Note 1 to entry: In accordance with intended use, clinical performance can include expected values, diagnostic sensitivity and diagnostic specificity based on the known clinical condition or physiological/pathological process/state of the individual, and negative and positive predictive values based on the prevalence of the disease.

[SOURCE: GHTF/SG5/N6:2012]

Clinical Performance Study

Study undertaken to establish or confirm the clinical performance of an IVD medical device

Notes: Testing performed pre-market that is not designed to address clinical performance of an IVD medical device is not considered a clinical performance study (e.g. customer feedback studies, external analytical performance studies, research studies).

[SOURCE: GHTF/SG5/N6: 2012]

Clinical Performance Study Protocol (CPSP)

Document that states the rationale, objectives, design, risk, proposed analysis, methodology, monitoring, conduct and record-keeping of the clinical performance study.

Note: The CPSP need not be a single document but a series of documents related and referenced to each other for the purpose of creating the CPSP.

[SOURCE: ISO 20916:2019].

Clinical Performance Study Report (CPSR)

Document describing the objectives design, execution, statistical analysis, results and conclusion(s) of a clinical performance study.

Note 1 to entry: Some elements of the clinical performance study report can be covered by stand-alone documents that are references in the clinical performance study report.

Note 2 to entry: The CPSR need not be a single document but a series of documents related and referenced to each other for the purpose of creating the CPSR.

[SOURCE: GHTF/SG5/N8:2012]

Endpoint

Principal (primary) or secondary indicator used in a clinical performance study to assess the performance of the IVD medical device.

Note 1 to entry: For example, endpoints can be statistical measures for performance or clinical events/outcomes.

[SOURCE: ISO 20916:2019, 3.17]

Ethics committee (EC)

Independent body whose responsibility it is to review clinical investigations in order to protect the rights, safety and well-being of human subjects participating in a clinical investigation.

[SOURCE: ISO 14155:—1), 3.24, modified — Note 1 to entry has been removed.]

Informed consent

Process by which an individual voluntarily confirms willingness to participate in a particular clinical performance study, after having been informed of all aspects of the study that are relevant for the decision to participate.

[SOURCE: ISO 14155:—1), 3.27, modified]

Note 1 to entry: For the purposes of this document, the permission is typically for providing specimens or participating in a clinical performance study.

Note 2 to entry: The informed consent document lists the risk(s) and benefit(s) to the subject, when applicable.

Note 3 to entry: The information provided can be broad in nature, allowing the specimen to be used for future undetermined studies, or the information can be specific to a particular study.

Intended Use/ Intended Purpose

Objective intent of the manufacturer regarding the use of a product, process or service as reflected in the specifications, instructions and information provided by the IVD manufacturer

Note 1 to entry: Intended use statements for IVD labelling can include two components: a description of the functionality of the IVD medical device (e.g. an immunochemical measurement procedure for the detection of analyte "x" in serum or plasma), and a statement of the intended medical use of the examination results.

[SOURCE: ISO 18113-1:2009, 3.31, modified]

Interventional Clinical Performance Study

Study in which test results obtained during the study influence patient management decisions and guide treatments

EXAMPLE Prospective clinical studies of companion diagnostic device.

[SOURCE: GHTF/SG5/N8:2012]

Investigator Brochure

Compilation of analytical and clinical performance data relevant to the clinical performance study

Note 1 to entry: The investigator brochure includes risk/benefit information of the IVD device under investigation and sampling procedures.

[SOURCE: ISO 20916:2019, 3.22]

IVD Medical Device

Refer to AHWP/WG2-WG1/F001:2016 *Definition of the Terms "Medical Device" and "In Vitro Diagnostic (IVD) Medical Device"*.

Leftover Specimen/Leftover Sample

Unadulterated remnants of human derived specimens collected as part of routine clinical practice and after all standard analysis has been performed.

Note 1 to entry: Such specimens/samples would be otherwise discarded as there is no remaining clinical need for them.

Note 2 to entry: This can include specimens collected for research or other purposes not connected to the clinical performance study in question.

[SOURCE: ISO 20916:2019, 3.25]

Specimen

Discrete portion of a body fluid or tissue taken for examination, study, or analysis of one or more quantities or characteristics to determine the character of the whole.

[SOURCE: ISO 18113-1:2009, 3.54, modified — "Specimen" is the first preferred terms, Notes to entry have been removed.]

Specimen Collection Procedure

All the steps involved in collecting a specimen from a human subject. This includes all preparatory steps, the actual collection and any after-treatment, and the disposal of any procedure-related materials.

EXAMPLE Fasting, pre-medication, anaesthesia procedures, blood draw, biopsy, disposal of sharps.

[SOURCE: ISO 20916:2019, 3.48].

Sponsor

Individual or organization taking responsibility and liability for the initiation, implementation and oversight of a clinical performance study.

Note 1 to entry: when an investigator initiates, conducts and takes full responsibility for a clinical performance study, the investigator also assumes the role of sponsor and is identified as the sponsor-investigator.

[SOURCE: ISO 20916:2019, 3.49]

Study Site

Institution(s) or location(s) where the clinical performance study is carried out, under the supervision of a principal investigator (3.32)

Note 1 to entry: For the purpose of this document, "study site" is synonymous with "study centre".

Note 2 to entry: Study sites include testing locations and specimen collection sites but not commercial procurement entities, e.g. providers of archived specimens (such as biobanks).

[SOURCE: ISO 20916:2019, 3.50]

Subject

Human who participates in a clinical performance study or whose specimen is used in the study

Note 1 to entry: Depending on the study, a subject can be either a healthy individual or a patient.

[SOURCE: ISO 20916:2019, 3.51]

5.0 Purpose of Clinical Performance Studies

The purpose of a clinical performance study is to establish or confirm aspects of IVD medical device performance which cannot be determined by analytical performance studies, literature and/or previous experience gained by routine diagnostic testing. This information is used to

demonstrate compliance with the relevant Essential Principles with respect to clinical performance. When a clinical performance study is conducted, the data obtained is used in the performance evaluation process and is part of the clinical evidence for the IVD medical device (see AHWP/WG5/F004:2015 – '*Clinical Evidence for IVD medical devices–Scientific Validity Determination and Performance Evaluation'*). These studies are typically undertaken in the pre-market phase of the IVD medical device to support its compliance with the EPs, but may also be undertaken in the post-market phase.

Detailed requirements for the conduct of a clinical performance study are set out in ISO 20916:2019 In vitro diagnostic medical devices — Clinical performance studies using specimens from human subjects — Good study practice.

6.0 Clinical Performance Study Design Type

Clinical performance studies should be designed in such a way as to maximize the relevance of the data while minimizing potential biases.

Designs of IVD medical device clinical performance studies are either observational or interventional. An observational study refers to a study in which test results obtained during the study are not used for patient management and do not impact treatment decisions. An interventional study refers to a study in which test results obtained during the study may influence patient management decisions and may be used to guide treatments.

Observational design can be further characterized as a combination of the following designs:

- **Cross-sectional design (single time-point design)** – testing of one or few specimens/samples per patient that are collected at a single time-point
- **Longitudinal design** – testing of multiple specimens/samples per patient that are collected over an extended period of time (e.g. weeks, months, years)
- **Retrospective design** – testing of previously collected specimens for which the analyte status and the patient's clinical status is known (characterized specimens) prior to the commencement of the study
- **Prospective design** – testing of specimens collected before or during the study but for which both the analyte status and the patient's clinical status are established during the study
- **Prospective-retrospective design** – testing of previously collected specimens for which the clinical status is known but the analyte status is unknown and will be established during the study

NOTE: The terminology for the various study designs can be used in the context of either diagnostic or epidemiologic studies, and these terms are associated with definitions that vary accordingly. The definitions and methodologies described in this document refer to diagnostic test designs, not epidemiological designs.

Figure 1 provides an illustration of the use of study design types for a single test purpose. Further clarifications on these study design types are presented below.

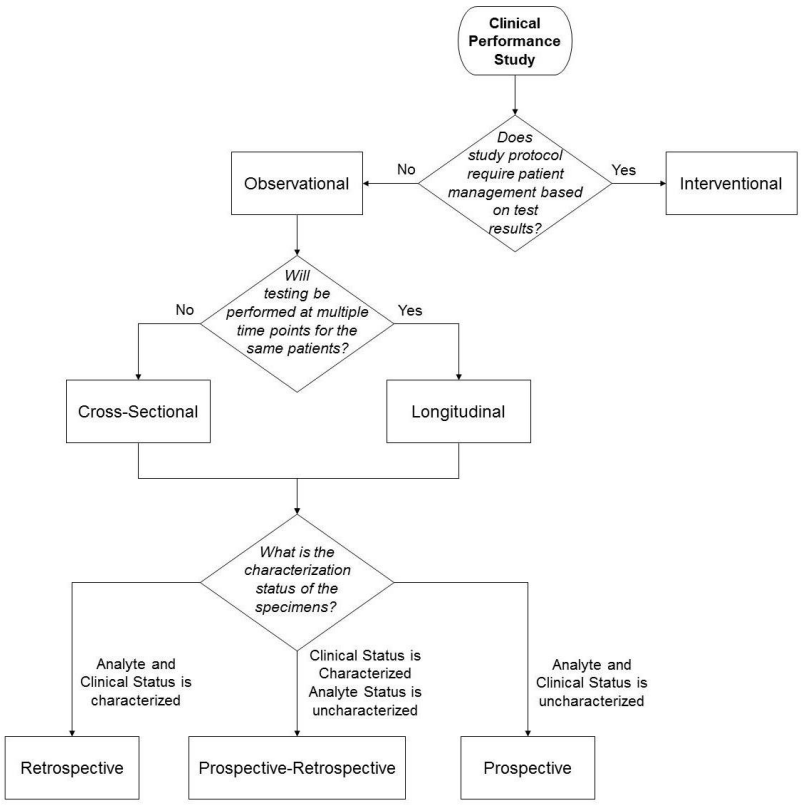

Figure 1: Decision tree for clinical performance study design

NOTE: In some cases it may be necessary to use a combination study design which includes both characterized and uncharacterized specimens.

6.1 Observational Designs

The majority of IVD medical device clinical performance studies follow an observational design. This design is common and appropriate because such study results are not used to determine patient management decisions as they are done in parallel to the routine diagnostic testing. The specimens for these studies are derived from:

- previously-collected specimens, such as archived specimens or leftover specimens that would otherwise have been discarded (also called surplus specimens); and/or
- additional specimens collected specifically for the study purposes

The following sections outline design types that further characterize studies of an observational nature.

6.1.1 Cross-sectional Designs

Cross-sectional clinical performance studies are those where correlation of test results to clinical condition can be established at a single time-point. Examples of cross-sectional clinical performance studies are presented below.

Example 1 (Diagnosis): Troponin for the diagnosis of myocardial infarction.

Example 2 (Screening): Detection of Human Papilloma Virus (HPV) in asymptomatic patients.

Example 3 (Predisposition): Detection of single nucleotide polymorphisms (SNPs) in particular genes (e.g. *HSPA1A*) to estimate the risk of developing cardiac disease.

Example 4 (Prognosis): Breast cancer gene expression profile test to determine the risk for distant metastasis.

Example 5 (Prediction): Detection of KRAS mutation as a marker of probable failure of epidermal growth factor receptor (EGFR)-targeted therapy.

NOTE: Refer to Appendix for a description of the common test purposes (e.g. screening, diagnosis, monitoring) for IVD medical devices.

In some cases, it is necessary to follow a modified version of cross-sectional design whereby testing is performed only at the initial time point, but patients are evaluated at later time points (follow-up) to determine their clinical status. This modification is known as a 'delayed cross-sectional design' and would apply in the following situations:

- the IVD medical device is used to evaluate the likelihood of future states (i.e. predisposition, prognosis and prediction); or
- there exists no applicable method to establish the clinical state at the time of testing or the method is deemed too invasive, thereby requiring follow-up to determine clinical outcome

6.1.2 Longitudinal Designs

Longitudinal clinical performance studies involve multiple patient measurements of the same analyte over time to validate the clinical performance of an IVD medical device. Examples of longitudinal clinical performance studies are presented below.

Example 1 (Diagnosis): In women less than 40 years old who have had amenorrhea for 4 months or more, serial FSH testing is useful for the diagnosis of primary ovarian insufficiency (i.e. two FSH measurements, obtained at least 1 month apart, in the menopausal range).

Example 2 (Screening): In patients at risk of renal failure, serial cystatin C measurements can be used to screen for early renal function decline.

Example 3 (Monitoring): In HIV infected patients the measurement of viral load after first establishing the baseline value can be used to assess treatment response.

Example 4 (Predisposition): Determination of the somatic mutation rate to estimate the risk of developing cancer.

Example 5 (Prognosis): In heart failure patients, a minimal change in B-type natriuretic peptide (BNP) levels upon initiation of therapy is associated with a low risk of mortality.

Example 6 (Prediction): For chronic myeloid leukemia (CML) patients undergoing imatinib treatment, significant decrease of BCR-ABL mRNA levels over time is predictive of continued treatment response.

6.1.3 Retrospective and Prospective Designs

This section describes clinical performance studies that follow a retrospective design, studies that follow a prospective design, studies that follow prospective-retrospective study design and studies that combine multiple designs.

Retrospective studies are appropriate if the following criteria are addressed:

- the specimens are representative of the intended use population (e.g. reflects variability of the clinical condition not just typical cases);

- the specimens are derived from a sufficiently large number of study subjects to reflect a random sampling;
- there is adequate data related to the clinical status of patients;
- the samples span the assay range (if applicable);
- there is minimal bias due to specimen/sample selection; and
- the analyte is stable over time.

When the criteria for retrospective study design eligibility are not met, the clinical performance study would need to follow a prospective design.

In the case of IVD medical devices used for the determination of a patient's future state (e.g. predisposition, prognosis, prediction), the clinical performance study will often be based on a prospective design. However, a retrospective design or a prospective-retrospective design could be used if specimen collection protocols were controlled to ensure that results are not biased or confounded.

A prospective-retrospective study employs a design that uses specimens that were collected previously under another protocol. These specimens will not be characterized for the analyte in question but will be clinically characterized. The specimens are collected retrospectively and are analytically characterized prospectively during the study.

Randomized controlled trials for therapeutic agents present a potential valuable source of specimens for a prospective-retrospective study as the patient outcome is known. Use of these specimens for prognosis or prediction is subject to the following considerations:

- specimens were collected prior to treatment; and
- specimens for testing reflect a relatively equal distribution from each of the study arms (treatment groups)

For certain clinical conditions (e.g. low disease prevalence), it may be necessary to design a clinical performance study that combines retrospective and prospective study designs. For example, a clinical performance study of an HIV-1/2 assay intended for donor screening may involve random blood donors (i.e. prospective uncharacterized specimens), HIV/AIDS patients (i.e. retrospective specimens characterized according to clinical status), and HIV-2 antibody-positive specimens (i.e. retrospective specimens characterized according to analyte status).

6.2 Interventional Designs

If performance claims for an IVD medical device cannot be demonstrated by an observational clinical performance study design, an interventional design would be appropriate when:

- there exists no established method for making decisions on patient management and the use of archived specimens would not be suitable to demonstrate the intended performance claims; or
- the manufacturer intends to demonstrate that the use of the IVD medical device impacts patient clinical outcome; or
- an IVD medical device is co-developed alongside a therapeutic product such that the information provided by the IVD medical device will influence the patient treatment in a therapeutic clinical trial (e.g. stratification of treatment arm).

This design uses specimens specifically collected for the study and the results of the clinical performance study would be used in patient management decisions.

The following examples outline situations when an interventional design would be required:

Example 1 (Diagnosis): To determine if patient outcome (i.e. treatment efficacy) is improved by correct infectious disease identification (e.g. Hepatitis A versus Hepatitis B) thereby facilitating selection of optimal treatment regimens.

Example 2 (Screening): To determine if prenatal screening for a particular genetic developmental disorder improves patient outcome (i.e. disease mitigation) because treatment can begin immediately following birth.

Example 3 (Monitoring): To determine if patient outcome (i.e. treatment efficacy) is improved by regular monitoring of changes to analyte concentration.

Example 4 (Predisposition): To determine if prophylactic interventions and/or lifestyle changes improve outcomes for patients at high risk of developing late-onset genetic conditions.

Example 5 (Prognosis): To determine if patient outcome (i.e. treatment efficacy) is improved by accurate disease staging and more aggressive treatments for patients with worse prognosis.

Example 6 (Prediction): To determine if a marker predicts a differential efficacy or safety of a particular therapy based on the marker's status (e.g. to test if patients expressing the marker respond to the specific treatment or respond to a greater degree than those without the marker or that some will respond negatively on the treatment)

7.0 Clinical Performance Study Design Considerations

The design of a clinical performance study should provide the data necessary to address the clinical performance of the IVD medical device. The clinical performance study design should account for potential risks, should follow appropriate ethical principles, and should be compliant with all relevant legal and regulatory requirements.

The choice of the design for the clinical performance study will depend on the following considerations:

- study objectives;
- intended use, specifically:
    - test purpose(s) (e.g. diagnosis, screening, monitoring; Refer to Section 7.1 and Appendix);
    - target population(s) (e.g. age, race, gender, geography, clinical condition, treatment status);
    - specimen type(s) (e.g. serum, plasma, urine, whole blood);
    - intended user(s)/operator(s) (person performing the test e.g. lay person, medical profession/lab technician).
- quality, availability and accessibility of specimens (e.g. limited number of leftover specimens available);
- intended use setting's environmental conditions;
- established analytical performance characteristics (e.g. precision, interference, measuring interval (range), cut-off, limit of detection, limit of quantification);
- intended clinical performance characteristics (e.g. sensitivity, specificity, positive predictive value, negative predictive value, reference intervals, cut-off);
- prevalence of the clinical condition/physiological or pathological state;
- novelty of the technology and/or clinical use (e.g. relevant previous experience);

- availability of appropriate method(s) to establish the clinical status of the subject[1];
- availability of quality control material;
- mechanisms to avoid bias.

The following additional considerations will also need to be taken into account in designing the clinical performance study

- specimen/sample collection;
- handling and storage conditions (e.g. sample cannot be frozen) (Refer to Section 7.2);
- clinical performance study site (e.g. point-of-care setting, central laboratory, testing site (Refer to Section 7.3);
- sample size estimate;
- description of planned statistical analysis (Refer to Section 7.4);
- the outcome of the risk evaluation; (Refer to Section 7.5);
- ethical considerations (Refer to Section 7.6).

7.1 Test purpose

An IVD medical device may be designed for a variety of intended uses with different test purposes (e.g. diagnosis, screening, monitoring, predisposition, prognosis, prediction). See Appendix for details.

The test purpose will directly influence the subject sample size (N) and selection criteria (including inclusion and exclusion) when planning and designing the clinical performance study. For example, if the disease state prevalence is low and the intent of the assay is to screen asymptomatic individuals, specimens from a large number of subjects may be required to provide sufficient evidence of clinical performance. However, if the assay is to be used for diagnosis in symptomatic individuals, specimens from a smaller number of subjects may be adequate.

Where appropriate, there should be consideration regarding the timing of specimen collection such as prior to treatment or during treatment. For example, a test for predisposition would require that the specimens should be drawn prior to the onset of the condition.

Where appropriate, the study should be designed to include patient follow-up to determine their clinical endpoint/outcome. This would be applicable for tests that identify future conditions such as tests for predisposition, prognosis and prediction.

Where appropriate, multiple test purposes might be evaluated simultaneously. In these cases more than one design type can be combined into a single clinical performance study. Such clinical performance studies should be designed and involve patient populations (with known or readily identifiable clinical status) that would sufficiently validate all of the potential test purposes.

7.2 Specimen Collection and Handling

Samples used in clinical performance studies are derived from specimens which may be obtained from several sources, including purposefully-collected specimens, leftover specimens, or archived specimens.

1 This is a critical point. Any reference standard used needs to clearly support the clinical function of the IVD and have a published and accepted strong body of clinical evidence. Discussion on the role of reference standards and examples of when and how they may be used are found in documents listed in the 3.0 References.

Purposefully-collected specimens refer to specimens that were drawn from patients with the specific intention of using them in a particular clinical performance study. These specimens or their derived samples may be tested immediately after collection (i.e. fresh) or may be stored (e.g. refrigerated or frozen) for testing at a later date.

Information about the specimens, when applicable and subjects providing specimens should be recorded:

- Validated specimen type (for example only plasma collected using validated anticoagulant);
- Inclusion criteria;
- Exclusion criteria;
- Information necessary to characterise the subject/specimen (e.g. status of other analytes, concomitant medications);
- Number of specimens and/or subjects;
- Specimen storage, handling, transport, and disposal.

Leftover specimens are considered to be remnants of specimens collected for routine diagnostic testing that would otherwise have been discarded, or specimens that were previously collected for other research purposes (e.g. basic research studies, pharmaceutical clinical trials, previous IVD medical device clinical performance studies).

Archived specimens or samples refer to specimens or samples that were collected in the past and are obtained from repositories (e.g. tissue banks, commercial vendor collections). While archived specimens or samples tend to be well-characterized whereby the analyte status and/or clinical status of the patient are known, some specimens or samples may be archived without first establishing their analyte/clinical status. Well-characterized archived specimens or samples are often the source for unique or rare specimens or samples in sufficient quantity, whereby without these specimens or samples, it would be difficult, if not impossible, to conduct the study in a reasonable timeframe.

For many IVD medical device clinical performance studies, it is appropriate to use leftover or archived specimens in lieu of purposefully-collected specimens provided that sufficient information concerning specimen characterization is available (e.g. anticoagulant, known patient treatment(s) that may influence test results, time of last treatment dose for therapeutic drug monitoring assays). Generally, clinical performance characteristics can be validated using leftover or archived specimens if they were collected, handled and stored appropriately (i.e. good integrity) and if the information available for the specimens meets the study design requirements. Care must be taken to avoid introducing a selection bias through the use of leftover or archived specimens. Documentation should exist to support the integrity of the selected specimens in the clinical performance studies.

Attention needs to be given when using leftover or archived specimens to evaluate certain prognostic or predictive tests due to the potential that treatment regimens may bias results. It is generally more appropriate to use leftover or archived specimens when there is minimal treatment heterogeneity among the patient populations. Such an approach reduces the risk that variable treatment regimens will obscure the influence of prognostic or predictive markers on clinical outcomes.

7.3 Clinical Performance Study Site Location

Clinical performance studies are generally performed at sites external to the manufacturer although the manufacturer’s site can be one of the testing sites included in the study. However,

in certain circumstances it might be appropriate to perform the testing only at the manufacturer's site; in this case a justification for this decision should be provided. For example, a study to determine expected values can often be performed entirely at the manufacturer's site.

Sites are often chosen to serve as specimen collection sites only and do not perform testing. Similarly, sites can perform testing using the investigational IVD medical device but do not have the ability to perform testing using the reference method. In these cases, extra care is needed to minimize bias between study sites.

It is important to ensure that the collection and testing sites are reflective of the intended use environment and/or intended user.

Individual study sites need not be identified in the study protocol, however, the sponsor should maintain an updated list of principal investigators, study sites, and institutions. This list can be kept separately from the CPSP. The final list should be provided with the clinical performance study report.

7.4 Statistical Design

In designing the study, statistical considerations should be specified in advance and be based on sound scientific principles and methodology. Care must be taken in developing a statistical plan that includes consideration of, for example, the following:

- statistical design, method and analytical procedures;
- level of significance and power of the clinical performance study;
- appropriate sample size (N) for estimation of clinical performance measures (e.g. sensitivity and specificity) with confidence intervals;
- appropriate subject inclusion and exclusion criteria (e.g. age, disease status);
- appropriate specimen/sample inclusion and exclusion criteria (e.g. specimen/sample integrity);
- minimization of bias such as selection bias, spectrum bias, verification bias (e.g. specimen/sample selection, collection, handling and storage; blinding of operator to clinical status of patient);
- criteria for re-examination and resolution (e.g. equivocal results, discrepant results);
- criteria for data exclusion (e.g. protocol deviation);
- analysis methodology;
- clinically relevant performance characteristics (e.g. sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, positive predictive value, negative predictive value, prevalence, percent agreement, correlation to clinical endpoints/outcomes, expected values);
- pass/fail criteria to be applied to the results of the clinical performance study;
- provision for an interim analysis, when applicable;
- procedures that ensure that all the data is taken into account; and
- treatment of missing, unused or spurious data.

7.5 Potential Risks

A variety of factors influence the potential risks to patients when conducting clinical performance studies for IVD medical devices.

When study specimens are obtained from specimens already taken for routine diagnostic testing there is no additional risk coming from the study. However, if the specimen is collected specifically for the study and involves invasive collection procedures the risks associated with these procedures should be taken into consideration. In doing so the level of invasiveness of the sampling procedures should also be taken into account (e.g. venepuncture versus spinal puncture).

Interventional studies carry higher risks as the results are used to manage patients. For these studies, appropriate procedures for adverse events monitoring and handling should be in place.

For clinical performance studies in which there is a risk to patients, clinical performance studies should only be performed once the analytical performance of the device has been established and determined to be acceptable. For example, in clinical performance studies for companion diagnostics, the way in which the biomarker status impacts treatment decisions may pose a risk to the patients.

7.6 Ethical Considerations for Clinical Performance Studies

As a general principle, the rights, safety and well-being of subjects participating in IVD medical device clinical performance studies shall be protected in accordance with the ethical principles laid down in the Declaration of Helsinki.

It is ethically important in deciding to conduct a clinical performance study that it should generate new data and answer specific safety and/or performance questions that remain unanswered by the current body of knowledge. The desire to protect human subjects from unnecessary or inappropriate experimentation must be balanced with the need to protect public health through the use of clinical performance studies where they are indicated based on scientific needs (e.g. specific mutations in a given population). In all cases, however, care must be taken to ensure that the necessary data are obtained through a scientific and ethical manner that does not expose subjects to undue risks or discomfort. The rights, safety and well-being of subjects are paramount, and appropriate clinical performance study design and conduct is essential to generate meaningful data.

7.6.1 Informed Consent

The requirement for patient informed consent should be based on the risk posed to subjects participating in the clinical performance study. For IVD medical devices, informed consent is required for the following types of clinical performance studies:

a) an interventional clinical performance study, or

b) a study in which specimen collection is done primarily for the purpose of the clinical performance study and where the specimen collection procedures pose additional risks to the subject, or

c) when the conduct of the study involves additional risks for the subjects of the studies.

If the clinical performance study will solely use leftover or archived specimens that are not individually identifiable (i.e. devoid of information that would otherwise permit traceability to the patient of origin), the requirement for informed consent may not apply or may be waived.

In some cases, informed consent may exist at the participating site in a general form to cover the use of the specimens in any clinical performance studies.

The need for informed consent should be discussed with the ethics committee of each institution included in the study. The process for obtaining informed consent should be documented, including the process for providing subjects with new information, as needed, and

if necessary, the process for obtaining informed consent in circumstances when the subject is unable to physically provide a signature.

7.6.2 Ethics Committee Involvement

Prior to commencing a clinical performance study, the participating sites may require approval by their local ethics committee. This independent committee is formally designated to review, approve and monitor studies involving human subjects with the aim of protecting the rights, safety and well-being of the subject. Where applicable, the principle investigator is required to:

- provide the sponsor with copies of any clinical performance study related communications between the principal investigator and the ethics committee, when they exist;
- obtain the written and dated approval/favourable opinion, or a waiver of the ethics committee for the clinical performance study before initiating and implementing all subsequent amendments;

  NOTE In certain jurisdictions, no response from the ethics committee after a defined time limit constitutes a waiver.
- perform adverse event recording and reporting.

Subject to national regulations, the communication with the ethics committee can be performed by the sponsor, partly or in full, in which case the sponsor shall keep the principal investigator informed. The ethics committee may choose to exempt certain IVD medical device clinical performance studies from their approval.

7.6.3 Communicating Test Results Outside of the Study

In some rare instances, there may be a need to communicate clinical performance study results to clinicians and/or public health institutions during or after the study. This can be the case when test results may have an immediate health impact (e.g. no routine testing exists) to the patient, patient's relatives or the public.

The decision and the mechanism to report results to clinicians and or public health institutions should be discussed with the local ethics committee prior to beginning the clinical performance study.

> Example: An IVD medical device for the detection of SARS is undergoing a clinical performance study. Specificity is assessed using prospectively-collected patient specimens. During testing of the specificity population, a sample is reported as positive. This result is confirmed by testing with an alternate method. It is ethically appropriate to communicate such unexpected results to ensure appropriate patient treatment.

8.0 Clinical Performance Study Protocol

The clinical performance study protocol sets out how the study is intended to be conducted. It contains important information about the study design such as the identification and description of the IVD medical device under investigation, overall synopsis, sponsor, purpose, objectives, study population, description of test method(s) and interpretation of results, study site(s) training and monitoring, specimen type, specimen collection, preparation, handling and storage, inclusion and exclusion criteria, limitations, warning and precautions, data collection/management, data analysis, required materials, number of study sites and if applicable, clinical endpoints/outcomes, statement of conformity, informed consent process,

adverse events/adverse device effects/device deficiencies and requirements for patient follow-up.

In addition, the clinical performance study protocol identifies the key factors which may impact the completeness and significance of results, such as intended participant follow-up procedures, decision algorithms, discrepancy resolution process, masking/blinding, approaches to statistical analyses, and methods for recording endpoints/outcomes and, where appropriate, communication of test results.

Discussion with regulatory authorities or a conformity assessment body may be appropriate when there is uncertainty as to whether the proposed clinical performance study protocol is adequate.

9.0 Conduct of Clinical Performance Studies

A properly conducted clinical performance study, including compliance to the clinical performance study protocol and local laws and regulations (e.g. pre-study authorization for interventional studies by the RA), ensures the protection of subjects, the integrity of the data and that the data obtained is acceptable for the purpose of demonstrating conformity to the relevant Essential Principles of Safety and Performance.

Considerations for the conduct of a clinical performance study may include:

- Responsibility of the sponsor;
- independence of study personnel;
- qualification and training of study personnel;
- adequate infrastructure;
- appropriate calibration procedures and means of control;
- relevant method for determining the true clinical status of patient specimens
- (routine) monitoring of the study site;
- close-out activities;
- reporting of the monitoring activities;
- security and confidentiality of data.

Raw data of a clinical performance study should be maintained based on the Quality Management System (QMS) requirements.

10.0 Clinical Performance Study Report

The protocol, results and conclusions of a clinical performance study should be documented in a Clinical Performance Study report. The results and conclusions should be transparent, free of bias, and clinically relevant. The report should contain sufficient information to enable it to be understood by an independent party without reference to other documents.

Such a report should also include as appropriate any protocol amendments or deviations, and data exclusions with the appropriate rationale.

The level of detail of the clinical performance study report will vary based on the class of the IVD medical device:

- Class A – A clinical performance study would not be expected hence a report would not be required.
- Class B – The complete study report includes the method of data analysis, the study conclusion and relevant details of the study protocol.

- Class C – The complete study report includes the method of data analysis, the study conclusion and relevant details of the study protocol.
- Class D – The complete study report includes the method of data analysis, the study conclusion, relevant details of the study protocol, and typically the individual data points (formatted raw data).

11.0 Adverse events, adverse device effects and device deficiencies

When the clinical performance study uses leftover/archived specimens, subjects will not be at risk from adverse events. However, when the study is one in which the specimen is taken for the purposes of the study, it is possible that there might be adverse events impacting the subjects or user/operator may experiences an adverse device effect during the specimen collection procedures. In such a case, clinical safety reporting must be implemented by the principal investigator or the sponsor in accordance with the ethics committee/national requirements.

Appendix

The table below lists the most common test purposes and provides examples to illustrate their differences.

TABLE 1: DESCRIPTION OF COMMON TEST PURPOSES FOR IVD MEDICAL DEVICES

Test Purpose	Description	Examples
Diagnosis	Diagnostic tests are used to determine, verify or confirm a patient's clinical condition as a sole determinant. This type of testing also includes sole confirmatory assays (to verify results of previous testing) and sole exclusion assays (to rule out a particular condition). These tests are designed to evaluate a patient's current state.	▪ genetic test for the diagnosis of Tay-Sachs ▪ HBs antigen confirmatory assay to verify positive screening results ▪ D-dimer assay for exclusion of deep vein thrombosis ▪ karyotype testing for the diagnosis of Trisomy 18 (Edward's syndrome)
Aid to Diagnosis	Aid to Diagnosis tests are used to provide additional information to assist in the determination or verification of a patient's clinical status. The test is not the sole determinant. These tests are designed to evaluate a patient's current state.	▪ troponin test as an aid in myocardial infarction diagnosis ▪ genetic testing to aid in the diagnosis of familial hypercholesterolaemia (FH) ▪ thyroid-stimulating hormone test to evaluate thyroid function ▪ toxoplasma IgG avidity assay to determine likelihood of active infection ▪ ANA test for autoimmune disease determination ▪ test for genotyping of the Factor V Leiden mutation as an aid to diagnosis of thrombophilia
Screening	Screening tests are used to determine the status of a disease, disorder or other physiological state in an asymptomatic individual. These types of tests include genetic screening assays, tests for physiological typing, and tests used to reduce the risk of infectious disease transmission, such as assays for prenatal screening and donor screening (transfusion or transplantation). Depending on the nature of the condition and the targeted patient population, screening tests may be used routinely or may be restricted to 'at risk' patients. These tests are designed to evaluate an individual's current state.	▪ test to detect hepatitis B surface antigen in donated blood ▪ prenatal rubella IgG screening in pregnant women ▪ prenatal genetic testing for trisomy 21 (Down's syndrome) ▪ newborn genetic testing for phenylketonuria ▪ tests for the determination of HLA, blood groups and blood group factors for donor matching

TABLE 1: DESCRIPTION OF COMMON TEST PURPOSES FOR IVD MEDICAL DEVICES

Test Purpose	Description	Examples
Monitoring	Monitoring tests are used for the measurement of analyte levels for the purpose of adjusting treatments/interventions as required. Monitoring tests include the following: ▪ Assays which are used to ensure that an analyte remains within physiological levels or within an established therapeutic drug range. These types of monitoring tests are designed to evaluate an individual's current state. ▪ Assays which are used for serial measurement, whereby multiple determinations are taken over time. These types of monitoring tests are typically used for the detection/assessment of disease progression/regression, disease recurrence, minimum residual disease, response/resistance to therapy, and/or adverse effects due to therapy. These types of monitoring tests are designed to evaluate changes in an individual's state	▪ intraoperative iPTH monitoring during parathyroidectomy surgery to confirm removal of abnormal tissue ▪ self-test glucose monitoring to allow for quick responses to hyperglycemia or hypoglycemia ▪ therapeutic drug monitoring of immunosuppressants to prevent rejection of transplanted organs ▪ viral load testing of patients known to be infected with HIV to determine treatment response and adjust therapy if necessary ▪ monitoring of CA 15-3 levels in breast cancer patients in remission to detect recurrence ▪ test for the detection of BCR-ABL transcripts to monitor response/resistance in patients undergoing treatment for acute lymphoblastic leukemia (ALL) or chronic myeloid leukemia (CML) ▪ test for immunoglobulin and T-cell receptor gene rearrangements for the detection of minimal residual disease in cancer patients.
Predisposition	Predisposition assays are used to determine the likelihood of disease onset (i.e. assessing the risk of developing the disease in future) in presymptomatic patients. For patients at sufficient risk (as determined by test results), preventive interventions may be taken. These tests are designed to evaluate a patient's future state.	▪ genetic test for apolipoprotein E to assess the risk of developing Alzheimer's disease ▪ BRCA1/BRCA2 mutation status testing to assess the risk of developing breast cancer (patient may choose to have prophylactic mastectomy if they are at sufficient risk)
Prognosis	Prognostic tests are used to measure factors linked to clinical outcome irrespective of treatment. Such tests may be used to estimate the natural progression of a disease (i.e. outcome in the absence of treatment), or to determine the likelihood of a clinical outcome irrespective of therapeutic intervention. These tests are designed to evaluate a patient's future state.	▪ high sensitive C-reactive protein measurement for the risk stratification of patients with acute coronary syndromes to determine the likelihood of future cardiac events ▪ measurement of baseline HIV-1 RNA level to assess patient prognosis ▪ cancer gene expression profile testing for metastasis risk to tailor treatment aggressiveness

TABLE 1: DESCRIPTION OF COMMON TEST PURPOSES FOR IVD MEDICAL DEVICES

Test Purpose	Description	Examples
Prediction (of Treatment Response or Reaction)	Predictive tests are used to measure factors that determine the likelihood of patient responses or adverse reactions to a specific therapy. Predictive tests designed specifically for use with a targeted therapy are sometimes termed 'companion diagnostics' or 'personalized medicine'. These tests are designed to evaluate a patient's future state.	▪ HER-2/neu testing in breast cancer patients to assess likelihood of response to hormone therapy ▪ identification of variations in cytochrome P450 genes (i.e. metabolizer status) to determine potential therapeutic benefits and/or adverse reactions to antiplatelet treatment
Determination of Physiological Status	Physiological status determination tests are used to evaluate the physiological state of an individual for the purpose of identifying a human condition or characteristic. These tests are designed to evaluate a patient's current state.	▪ hCG test for the determination of pregnancy

Depending on its intended use, an IVD medical device may have one or more test purposes. For example, a nucleic acid-based infectious disease assay may be used for diagnosis (testing in patients suspected to be infected), screening (testing in asymptomatic patients) and monitoring (determination of viral load to assess effectiveness of treatment).

In some cases, it may be difficult to define a distinct test purpose, especially when one is dependent on (or linked to) another. For example, a single genetic test may be used to detect the genotype (i.e. screening) as well as providing the likelihood of developing the condition (i.e. predisposition).