



DPC: 25/30500985 DC

BSI Group Headquarters

389 Chiswick High Road London W4 4AL

Tel: +44 (0)20 8996 9000

Fax: +44 (0)20 8996 7400

www.bsigroup.com

Date: 09 May 2025

Origin: International

Latest date for receipt of comments: 25 June 2025

Project No. 2024/02767

Responsible committee: CH/62 Electrical Equipment in Medical Practice

Interested committees:

Title: Draft BS EN IEC 63521 ED1 BS EN 63521 ED1 Machine Learning-enabled Medical Device – Performance Evaluation Process.

Please notify admin.start@bsigroup.com if you are aware of any keywords that might assist in clarifying or identifying the standard or of the content of this standard:

- i) has any issues related to 3rd party IPR, patent or copyright
- ii) affects other national standard(s)
- iii) requires additional national guidance or information

WARNING: THIS IS A DRAFT AND MUST NOT BE REGARDED OR USED AS A BRITISH STANDARD. THIS DRAFT IS NOT CURRENT BEYOND 25 June 2025

This draft is issued to allow comments from interested parties; all comments will be given consideration prior to publication. No acknowledgement will normally be sent. **See overleaf for information on the submission of comments.**

No copying is allowed, in any form, without prior written permission from BSI except as permitted under the Copyright, Designs and Patent Act 1988 or for circulation within a nominating organization for briefing purposes. Electronic circulation is limited to dissemination by e-mail within such an organization by committee members.

Further copies of this draft may be purchased from BSI Shop <http://shop.bsigroup.com> or from BSI Customer Services, Tel: +44(0) 20 8996 9001 or email cservices@bsigroup.com. British, International and foreign standards are also available from BSI Customer Services.

Information on the co-operating organizations represented on the committees referenced above may be obtained from <https://standardsdevelopment.bsigroup.com/>.

E-mail: admin.start@bsigroup.com

Introduction

This draft standard is based on national and international discussions. Your comments on this draft are invited and will assist in the preparation of the consequent standard.

For international standards, comments will be reviewed by the relevant UK national committee before sending the consensus UK vote and comments to the international committee, which will then decide appropriate action. If the international standard is approved, it is usual for the text to be published as a British Standard.

For national standards, comments will be reviewed by the relevant UK national committee and the resulting standards published as a British Standard.

UK Vote

Please indicate whether you consider the UK should submit a negative (with supporting technical reasons) or positive vote on this draft. Please indicate if you are aware of any reason why this draft standard should not be published as a British Standard.

Submission of Comments

- **Annotated drafts are not acceptable and will be rejected.**
- All comments should be submitted online at <https://standardsdevelopment.bsigroup.com/>. You will need to register in order to comment.

Template for comments and secretariat observations

Date: xx/xx/20xx	Document: ISO/DIS xxxx
------------------	------------------------

1	2	(3)	4	5	(6)	7
MB	Clause No./ Subclause No./Annex (e.g. 3.1)	Paragraph/ Figure/Table/Note	Type of comment	Comment (justification for change) by the MB	Proposed change by the MB	Secretariat observations on each comment submitted
	3.1	Definition 1	ed	Definition is ambiguous and needs clarifying.	Amend to read '...so that the mains connector to which no connection...'	
	6.4	Paragraph 2	te	The use of the UV photometer as an alternative cannot be supported as serious problems have been encountered in its use in the UK.	Delete reference to UV photometer.	



62/545/CD

COMMITTEE DRAFT (CD)

PROJECT NUMBER:

IEC 63521 ED1

DATE OF CIRCULATION:

2025-05-02

CLOSING DATE FOR COMMENTS:

2025-07-25

SUPERSEDES DOCUMENTS:

62/474/NP, 62/486A/RVN

IEC TC 62 : MEDICAL EQUIPMENT, SOFTWARE, AND SYSTEMS

SECRETARIAT:

Germany

SECRETARY:

Ms Regina Geierhofer

OF INTEREST TO THE FOLLOWING COMMITTEES:

HORIZONTAL FUNCTION(S):

ASPECTS CONCERNED:

Safety

This document is still under study and subject to change. It should not be used for reference purposes.

Recipients of this document are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Recipients of this document are invited to submit, with their comments, notification of any local regulations or technical reasons that may exist and should be considered should this proposal proceed, recognizing that failure to address such requirements could result in the need for new or existing "In Some Countries" clauses. (SEE [AC/22/2007](#) OR [NEW GUIDANCE DOC](#)).

TITLE:

Machine learning-enabled medical device – Performance evaluation process

NOTE FROM TC/SC OFFICERS:

The relatively short comment period is motivated by the fact that at least one further Committee Draft is planned. This would allow for more timely comments and faster revision(s). The input of the P-members is urgently needed in order to align the standard as closely as possible with the needs of the medical device framework.

Copyright © 2025 International Electrotechnical Commission, IEC. All rights reserved. It is permitted to download this electronic file, to make a copy and to print out the content for the sole purpose of preparing National Committee positions. You may not copy or "mirror" the file or printed version of the document, or any part of it, for any other purpose without permission in writing from IEC.

AVA Documents
Document Sharing Hub

CONTENTS

FOREWORD.....	3
Introduction.....	5
1 Scope.....	6
2 Normative references	6
3 Terms and definitions	6
4 General requirements	11
5 MLMD performance evaluation	11
5.1 General principles of the performance evaluation framework	11
5.1.1 General	11
5.1.2 Scientific validity evaluation.....	14
5.1.3 Technical(analytical) performance evaluation	14
5.1.4 Clinical performance evaluation	15
5.1.5 Compilation of clinical evidence and submission.....	15
5.2 Three pillars of performance	16
5.2.1 General	16
5.2.2 Scientific validity - overview.....	16
5.2.3 Technical (analytical) performance - overview	17
5.2.4 Clinical performance - overview	17
5.3 Support for the evaluation	18
5.4 Monitoring for post-market	18
6 Evaluation process	18
6.1 MLMD performance evaluation planning	18
6.2 Scientific validity	18
6.2.1 General	18
6.2.2 Intended use.....	18
6.2.3 Is the scientific validity of the ML technology confirmed?	19
6.2.4 Verifications and identification of scientific validity	19
6.2.5 Provide evidence needed to assess scientific validity.	19
6.2.6 Appraisal and analysis of scientific validity information	19
6.3 Technical (analytical) performance evaluation.....	20
6.3.1 General	20
6.3.2 Planning and setup.....	20
6.3.3 Core evaluation & analysis	22
6.3.4 Advanced requirements	24
6.3.5 Reporting and improvement.....	25
6.4 Clinical performance evaluation	25
6.4.1 General	25
6.4.2 Clinical performance	25
6.4.3 Performance metrics.....	26
6.4.4 Feedback and continuous improvement	26
6.5 MLMD performance evaluation report	26
6.6 MLMD TPLC and 3-pillar evaluation	26
Annex A (informative) Relationship to other standards	28
A.1 General.....	28

Annex B (informative) MLMD and SaMD	29
Annex C (informative) MLMD Tasks	30
Annex D (informative) Performance metrics	31
D.1 General.....	31
D.2 Metrics.....	31
Annex E (informative) Hazards and performance evaluation	33
E.1 General.....	33
E.2 Examples of hazards and characteristics related to safety	33
E.3 Examples of events and circumstances	33
E.4 Examples of relationships between hazards, sequences of events, hazardous situations, possible harm and potential risk control measures	33
E.5 Example of relationship between characteristics related to safety and 3-pillar performance model	33
Annex F (informative) Evaluation report template	36
F.1 Section 1 - Identifications.....	36
F.2 Section 2 - Evaluation requirements.....	36
F.3 Section 3 - Evaluation specification.....	36
F.4 Section 4 - Evaluation methods.....	37
F.5 Section 5 - Evaluation Results	37
Annex G (informative) Comparison between MLMD product performance evaluation process and IEC 62304	38
Bibliography.....	40
Figure 1 – Relationship of the types of use	6
Figure 2 – Performance evaluation framework for MLMD	12
Figure 3 – Evaluation flow of the 3-pillar performance model	13
Figure 4 – Machine learning pipeline and mapping to the AI system life cycle	14
Figure 5 – Three pillars of MLMD performance	16
Figure 6 – Technical(Analytical) performance evaluation flow	20
Figure 7 – Example of 3-pillar performance evaluation in MLMD (3.11) TPLC	27
Figure A.1 – Relationship of key MEDICAL DEVICE standards to IEC 63521	28
Table D.1 –.....	31
Table E.1 – Example of relationship between characteristics related to safety and 3- pillar performance model	33
Table G.1 – Relationship to IEC 62304 ED1.....	38

INTERNATIONAL ELECTROTECHNICAL COMMISSION

Machine learning-enabled medical device –Performance evaluation process –

FOREWORD

- 1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as "IEC Publication(s)"). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
- 2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.
- 3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
- 4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.
- 5) IEC itself does not provide any attestation of conformity. Independent certification bodies provide conformity assessment services and, in some areas, access to IEC marks of conformity. IEC is not responsible for any services carried out by independent certification bodies.
- 6) All users should ensure that they have the latest edition of this publication.
- 7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.
- 8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.
- 9) IEC draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). IEC takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, IEC [had/had not] received notice of (a) patent(s), which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at <https://patents.iec.ch>. IEC shall not be held responsible for identifying any or all such patent rights.

IEC 63521 has been prepared by IEC Technical Committee 62: Medical equipment, software, and systems. It is an International Standard.

The text of this International Standard is based on the following documents:

Draft	Report on voting
XX/XX/FDIS	XX/XX/RVD

Full information on the voting for its approval can be found in the report on voting indicated in the above table.

The language used for the development of this International Standard is **English** [change language if necessary].

This document was drafted in accordance with ISO/IEC Directives, Part 2, and developed in accordance with ISO/IEC Directives, Part 1 and ISO/IEC Directives, IEC Supplement, available at www.iec.ch/members_experts/refdocs. The main document types developed by IEC are described in greater detail at www.iec.ch/publications.

The committee has decided that the contents of this document will remain unchanged until the stability date indicated on the IEC website under webstore.iec.ch in the data related to the specific document. At this date, the document will be

- reconfirmed,
- withdrawn, or
- revised.

Introduction

Generally, the concept of evaluation is the systematic determination of the extent to which an entity meets its specified criteria. The evaluation of product performance is vital to both the acquisition and development of any product. For MLMD, the relative importance of the various characteristics of product performance depends on the intended medical use of the product.

Performance evaluation of MLMD is more challenging than for SaMD because data-driven algorithms depend on amount and quality of the data as well as the model selection and training objective functions used. The **performance evaluation** of MLMD is considered more challenging than conventional medical device software.

Focusing on two key points, model and learning, the evaluation of ML performance is a crucial factor in measuring the overall aspect of MLMD performance. **Performance evaluation** of MLMD takes into account the continuous composition of various tasks during the lifecycle, such as ML modeling, ML verification and validation, deployment, and updating, MLMD products also need to be continuously evaluated in the total product lifecycle to validate whether relevant performance characteristics meet the quality requirements of the final product.

1 Scope

This document defines a standardized **performance evaluation** process for Machine Learning-enabled Medical Devices. The set of processes, activities, and tasks described in this document establishes a common framework for MLMD **performance evaluation** processes.

This process is to assist manufacturers to evaluate the **scientific validity**, the technical and clinical performance of the MLMD. It may be used for evaluating the performance to assure expected performance during post-market monitoring.

This document is applicable to the **performance evaluation** of all forms of MLMD, comprising ML components and, where appropriate, the integrated ML component with non-ML components. The most important factor to consider for **performance evaluation** is that ML provides the intended purpose and does not have unintended or unintentional effects on the safety and effectiveness of a medical device.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

- IEC Electropedia: available at <https://www.electropedia.org/>
- ISO Online browsing platform: available at <https://www.iso.org/obp>

3.1

abnormal use

conscious, intentional act or intentional omission of an act that is counter to or violates normal use and is also beyond any further reasonable means of user interface-related risk control by the manufacturer

EXAMPLE Reckless use or sabotage or intentional disregard of information for safety are such acts.

Note 1 to entry: See also IEC 62366-1, 4.1.3.

Note 2 to entry: An intended but erroneous action that is not abnormal use is considered a type of use error.

Note 3 to entry: Abnormal use does not relieve the manufacturer from considering non-user interface-related means of risk control.

Note 4 to entry: **Figure 1** shows the relationships of the types of use.

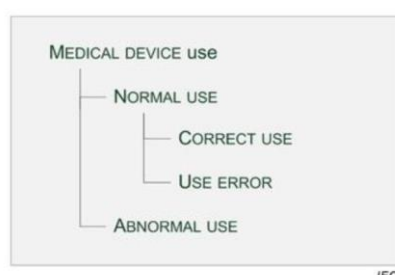


Figure 1 – Relationship of the types of use

Note 5 to entry: IEC 62366-1 Figure D.1 contains additional detail

[SOURCE: IEC 62366-1:2015, 3.1]

3.2

assessment

systematic process of collecting and analyzing data to determine the current status of target

Note 1 to entry: Adapted from ISO 10795:2019.

Note 2 to entry: Assessment target can be related to a product, a process, a system, a person or an organization

[SOURCE: ISO 10795:2019]

3.3

clinical evaluation

assessment of clinical data pertaining to a medical device to verify the clinical safety and performance of the medical device when used as intended by the [manufacturer](#)

[SOURCE: ISO 13485:2016, 3.3 modified]

3.4

clinical performance

behaviour of a [medical device \(3.34\)](#) and response of the [subject\(s\) \(3.50\)](#) to that medical device in relation to its intended use, when correctly applied to appropriate subject(s)

Note 1 to entry: Clinical performance can be defined under national regulations.

[SOURCE: ISO 14155:2020, 3.11]

3.5

clinical validation

process to establish the clinical evidence necessary to ensure that the intended purpose of the device is achieved in the target population

3.6

essential performance

performance of a clinical or diagnostic function where loss of the function or degradation of the performance beyond the limits specified by the [manufacturer](#) results in an unacceptable risk, other than that related to basic safety

Note 1 to entry: Clinical and diagnostic functions include, for example, therapeutic, patient monitoring and in vitro diagnostic functions

[SOURCE: IEC 60601-1:2005 and IEC 60601-1:2005/AMD1:2012, 3.27]

3.7

evaluation

systematic determination of the extent to which an entity meets its specified criteria

[SOURCE: ISO/IEC 12207:2008]

3.8

evaluation method

procedure describing actions to be performed by the evaluator in order to obtain results for the specified measurement applied to the specified product components or on the product as a whole

[SOURCE: ISO/IEC 25040:2011]

3.9

harm

injury or damage to the health of people, or damage to property or the environment

[SOURCE: ISO 14971:2019, 3.3]

3.10

hazard

potential source of [HARM](#)

[SOURCE: ISO 14971:2019, 3.4]

3.11

machine learning-enabled medical device

MLMD

medical device that uses machine learning, in part or in whole, to achieve its intended use

[Modified: IMDRF Machine Learning-enabled Medical Devices: Key Terms and Definitions (IMDRF/AIMD WG/N67), intended use]

3.12

manufacturer

natural or legal person with responsibility for the design and/or manufacture of a [medical device \(3.10\)](#) with the intention of making the [medical device \(3.10\)](#) available for use, under his name, whether or not such a [medical device \(3.10\)](#) is designed and/or manufactured by that person himself or on his behalf by another person(s)

[SOURCE: [ISO 14971:2019 \[1\]](#), 3.9]

3.13

measure, noun

variable to which a value is assigned as the result of measurement

Note 1 to entry: The term "measures" is used to refer collectively to base measures, derived measures, and indicators.

Note 2 to entry: Adapted from ISO/IEC 14598-1:1999.

[SOURCE: ISO/IEC 25040:2011]

3.14

measure, verb

ascertain or determine the magnitude or quantity of something

Note 1 to entry: make a measurement

[SOURCE: ISO/IEC 14598-1:1999]

[SOURCE: ISO 8000-2:2020, 3.4.1]

3.15

measurement

process of experimentally obtaining one or more values that can reasonably be attributed to a quantity

Note 1 to entry: set of operations having the object of determining a value of a measure.

[SOURCE: ISO/IEC 15939:2007]

3.16

misuse

usage in a way not intended by the [manufacturer](#) or the service provider

Note 1 to entry: Misuse includes human behaviour that is not intended but does not include deliberate system alterations or use of the system with the intention to cause harm.

Note 2 to entry: Misuse can result from overconfidence in the performance of the system.

Note 3 to entry: Depending on the causal relationship to the hazardous behaviour, there are two kinds of misuse, direct and indirect.

Note 4 to entry: Direct misuse, which could be a cause for the occurrence of a hazardous behaviour of the system, is considered to be a potential triggering condition (3.30). If its ability to contribute to the occurrence of hazardous behaviour is established, then it is considered to be a triggering condition. It is also possible that the direct misuse is part of a triggering condition, i.e. next to the direct misuse additional specific conditions of a scenario need to be present for the hazardous behaviour of the system to occur.

[SOURCE: ISO 21448:2022, 3.17]

3.17

misuse scenario

scenario in which misuse occurs

[SOURCE: ISO 21448:2022, 3.18]

3.18

MLMD product

set of [MLMD](#), procedures, and software, and associated documentation and data

3.19

normal use

operation, including routine inspection and adjustments by any user, and stand-by, according to the instructions for use or in accordance with generally accepted practice for those medical devices provided without instructions for use

Note 1 to entry: Normal use should not be confused with intended use. While both include the concept of use as intended by the manufacturer, intended use focuses on the medical purpose while normal use incorporates not only the medical purpose, but maintenance, transport, etc. as well.

Note 2 to entry: Use error can occur in normal use.

Note 3 to entry: Medical devices that can be used safely without instructions for use are exempted from having instructions for use by some authorities with jurisdiction.

Note 4 to entry: [Figure 1](#) shows the relationships of the types of use.

[SOURCE: IEC 62366-1:2015, 3.9]

3.20

performance

measurable result

Note 1 to entry: Performance can relate either to quantitative or qualitative findings

Note 2 to entry: Performance can relate to managing activities, process, products (including services), systems or organizations. [ISO/IEC 9000]

Note 3 to entry: characteristics defining the ability of a measuring instrument to achieve the intended functions. [IEC Electropedia]

[SOURCE: ISO/IEC 22989]

3.21

performance evaluation

systematic activities to determine measurable results against the set criteria

Note 1 to entry: Adapted from ISO 22300:2021

3.22

performance evaluation

assessment and analysis of data to establish or verify the [scientific validity](#), the technical(analytical) and, where applicable, the clinical performance of a [MLMD](#)

[SOURCE: IMDRF GRRP WG/N47: 2018, 3.21 modified]

3.23

process

system of activities, which uses resources to transform inputs into outputs

Note 1 to entry: Adapted from ISO 9000:2005.

[SOURCE: ISO/IEC 25040:2011]

3.24

risk

combination of the probability of occurrence of [HARM](#) and the severity of that [HARM](#)

[SOURCE: ISO 14971:2019, 3.18]

3.25

safety

freedom from unacceptable [RISK](#)

[SOURCE: ISO 14971:2019, 3.26]

3.26

scientific validity

demonstration of an association of ML technology (model, data, learning algorithm) to a clinical needs or medical purpose

3.27

task

<artificial intelligence>action required to achieve a specific goal

Note 1 to entry: Actions can be physical or cognitive. For instance, computing or creation of [predictions \(3.1.27\)](#), translations, synthetic data or artefacts or navigating through a physical space.

Note 2 to entry: Examples of tasks include classification, segmentation, prediction, regression, ranking, clustering and dimensionality reduction.

[SOURCE: ISO/IEC 22989:2022, 3.1.35]

3.28

technical performance

analytical performance

ability of a [MLMD](#) to correctly act a particular intended ML tasks (classification, detection, prediction, generation, etc.)

Note 1 to entry: measure of the results of functionalities or capabilities achieved for a project or programme during implementation (ISO/TR 21506:2018(en),3.82)"