



DPC: 24 / 30456154 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: 21 August 2024

Origin: International

Latest date for receipt of comments: 2 October 2024

Project No. 2022/01001

Responsible committee: CH/62 Medical equipment, software, and systems

Interested committees:

Title: BS EN IEC 63450 Testing of Artificial Intelligence / Machine Learning-enabled Medical Devices

Please notify the secretary if you are aware of any keywords that might assist in classifying or identifying the standard or if 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 2 October 2024**

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 <http://standardsdevelopment.bsigroup.com>

Responsible Committee Secretary: Ms Lindsey Ferrari (BSI)

Direct tel:

E-mail: lindsey.ferrari@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	Commend (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.	

**62/520/CD**

COMMITTEE DRAFT (CD)

PROJECT NUMBER:

IEC 63450 ED1

DATE OF CIRCULATION:

2024-08-16

CLOSING DATE FOR COMMENTS:

2024-11-01

SUPERSEDES DOCUMENTS:

62/411/NP, 62/418/RVN

IEC TC 62 : MEDICAL EQUIPMENT, SOFTWARE, AND SYSTEMS

SECRETARIAT:

Germany

SECRETARY:

Ms Regina Geierhofer

OF INTEREST TO THE FOLLOWING COMMITTEES:

SC 62A, SC 62B, SC 62C, SC 62D

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:

Testing of Artificial Intelligence / Machine Learning-enabled medical devices

NOTE FROM TC/SC OFFICERS:

Copyright © 2024 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.

CONTENTS

FOREWORD.....	4
INTRODUCTION.....	6
1 Scope.....	7
2 Normative references	7
3 Terms and definitions	7
3.22 Terms specific to Annex A	10
4 Test planning.....	10
4.1 Test plan.....	10
4.2 Test objectives.....	11
4.3 Test scope	11
4.4 Responsibilities and roles	11
4.5 Test environment	12
4.5.1 General	12
4.5.2 Test tools	12
4.5.3 Test data	12
4.6 Deliverables.....	12
4.7 Re-testing after changes	12
4.8 Test methods.....	13
4.8.1 General	13
4.8.2 Test oracle	13
4.8.3 Cross-Validation testing.....	13
4.8.4 Black-box test methods	19
4.8.5 White-box test methods	23
4.8.6 AI explainability/transparency	24
4.9 Statistics.....	26
4.9.1 General	26
4.9.2 Classification metrics.....	26
4.9.3 Regression metrics	26
4.10 Test phases / stages.....	27
5 Test specification	27
5.1 Test identification	27
5.2 Test conditions	27
5.3 Acceptance criteria	28
5.4 Test data	28
5.4.1 Test data selection	28
5.4.2 Data relevance	28
5.4.3 Data representativeness	28
5.4.4 Test data quantification	28
5.4.5 Test data preparation	28
5.4.6 DOUP (Data Of Unknown Provenance).....	29
5.4.7 Artificial data	29
5.4.8 Test data labelling	29
5.4.9 Use of unlabelled data	30
5.4.10 Traceability.....	30
6 Test results	30

6.1	Test results documentation	30
6.2	Test deviations	30
Annex A	(normative) Continuous learning	31
A.1	Scope	31
A.2	General concepts and requirements	31
A.3	Control mechanisms	32
A.3.1	Pre-testing by the manufacturer	32
A.3.2	Evaluation of training algorithm responsiveness	32
A.3.3	Limited learning	33
A.3.4	Evaluation of change	33
A.3.5	Utilization of back-to-back testing	33
A.3.6	Released model as a test oracle	33
A.4	New data samples	33
A.4.1	Training data	33
A.4.2	Test data	33
Annex B	(informative) Data/information leakage	35
B.1	Risks of data/information leakage and impact on resulting AI systems	35
B.2	Sources of data/information leakage	35
B.3	The ideal test set	35
B.4	How to avoid data/information leakage	35
Bibliography		37
Figure 1	– HCV Representation	14
Figure 2	– k-Fold CV Representation	15
Figure 3	– Repeated k-folds CV Representation	16
Figure 4	– Stratified CV by Age as a Defining Characteristic of the Population	17
Figure 5	– Time Series Cross-Validation Representation	18
Figure A.1	– Model release (3.22.1) and verification cycle	31

INTERNATIONAL ELECTROTECHNICAL COMMISSION

**TESTING OF ARTIFICIAL INTELLIGENCE / MACHINE LEARNING-ENABLED
MEDICAL DEVICES**

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 not 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> and www.iso.org/patents. IEC shall not be held responsible for identifying any or all such patent rights.

IEC 63450 was prepared by IEC technical committee TC62: 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

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.

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/standardsdev/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,
- replaced by a revised edition, or
- amended.

INTRODUCTION

In the rapidly evolving landscape of healthcare technology, artificial intelligence (AI) and machine learning (ML) are playing pivotal roles in [medical device](#) development. These technologies have the potential to enhance patient outcomes, improve diagnostics, and optimize treatment strategies. However, ensuring their safety, effectiveness, security, and transparency is critical.

To this end, the verification of medical devices through testing plays a crucial role, particularly when they are programmable. When dealing with AI/ML-enabled Medical Devices ([AI/ML-MD](#)), testing becomes more intricate due to their dynamic nature and reliance on complex algorithms, necessitating consideration of additional aspects.

As the main stakeholder responsible for verification and validation, this document addresses the manufacturer of medical devices. Nonetheless, the document can be used by any stakeholder that needs to perform testing of [AI/ML-MD](#). For example, a test laboratory can apply the requirements of this document by substituting each occurrence of “the manufacturer” by “the test laboratory”.

The document is organized in three main chapters, corresponding to the main testing activities of test planning, [test specification](#) and executing the test to get test results. An extended subsection of the test planning chapter is dedicated to explaining and setting the requirements for different test methods that can be used with [AI/ML-MD](#). The explanations are well separated from the requirements and are intended as only informative.

Two annexes complete the document. [Annex A](#) is normative and provides requirements for the case of [AI/ML-MD](#) that learn continuously in the field. Informative [Annex B](#) explains the problem of data/information leakage and provides some possible countermeasures.

For the purposes of this document:

- “shall” means that conformance with a requirement is mandatory for conformance with this document;
- “should” means that conformance with a requirement is recommended but is not mandatory for conformance with this document;
- “may” is used to describe a permissible way to achieve conformance with a requirement;
- “establish” means to define, document, and implement; and
- where this document uses the term “as appropriate” in conjunction with a required process, activity, task or output, the intention is that the manufacturer shall use the process, activity, task or output unless the manufacturer can document a justification for not so doing.

TESTING OF ARTIFICIAL INTELLIGENCE / MACHINE LEARNING-ENABLED MEDICAL DEVICES

1 Scope

Software testing is well treated in standardization (see [ISO/IEC/IEEE 29119-1:2022 \[1\]](#)). For medical devices, requirements on software testing are well established internationally by [IEC 62304:2006/Amd 1:2015 \[2\]](#). This document establishes additional aspects to specifically test [AI component](#) in [AI/ML-MD](#).

[AI component](#) considered in this document include locked [AI model](#) and [AI model](#) that learn under the control of the manufacturer. The case where the [AI model](#) is continuously learning in the field under indirect control by the manufacturer is treated in [Annex A](#).

Tests can be conducted at various lifecycle stages of the [AI/ML-MD](#), depending on the nature of the [AI model](#).

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC TS 4213:2022, *Information technology — Artificial intelligence — Assessment of machine learning classification performance*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC TS 4213:2022 , as well as the following apply.

ISO and IEC maintain terminological 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

A/B testing

statistical testing approach that allows testers to determine which of two systems or components performs better

[SOURCE: [ISO/IEC TR 29119-11:2020 \[3\]](#), definition 3.1.1]

3.2

AI component

device component based on AI technology

3.3

AI model

software that analyses datasets to find patterns and make predictions

3.4**AI/ML-MD****AI/ML-enabled medical device**

[medical device \(3.13\)](#) that has at least one [AI component](#)

3.5**AluT****Artificial Intelligence under Test**

instance of [AI component](#) that is the object of a [test \(3.17\)](#)

3.6**artificial data**

[synthetic data \(3.16\)](#) or [augmented data](#)

3.7**augmented data**

artificially generated, slightly modified copy of existing data

3.8**back-to-back testing****differential testing**

approach to testing whereby an alternative version of the system is used as a [pseudo-oracle \(3.15\)](#) to generate expected results for comparison from the same test inputs

EXAMPLE The [pseudo-oracle \(3.15\)](#) may be a system that already exists, a system developed by an independent team or a system implemented using a different programming language.

[SOURCE: [ISO/IEC TR 29119-11:2020 \[3\]](#), definition 3.1.16]

3.9**combinatorial testing**

black-box test design technique in which [test cases \(3.20\)](#) are designed to execute specific combinations of values of several parameters

EXAMPLE Pairwise testing, all combinations testing, each choice testing, base choice testing

[SOURCE: [ISO/IEC TR 29119-11:2020 \[3\]](#), definition 3.1.23]

3.10**DOUP****Data Of Unknown Provenance**

dataset that is already collected and generally available and that has not been collected for the purpose of being used for development or [test \(3.17\)](#) of an [AI/ML-enabled medical device](#) or dataset previously collected for which adequate records of the collecting process are not available

3.11**explainability**

property of an AI system to express important factors influencing the AI system results in a way that humans can understand

Note 1 to entry: It is intended to answer the question “Why?” without actually attempting to argue that the course of action that was taken was necessarily optimal.

[SOURCE: [ISO/IEC 22989:2022 \[4\]](#), definition 3.5.7]

3.12**extended back-to-back testing**

approach to testing whereby an alternative version of the system is used as a [pseudo-oracle \(3.15\)](#) to generate expected results for comparison

EXAMPLE The [pseudo-oracle \(3.15\)](#) may be a system that already exists, but uses different input data

Note 1 to entry: Difference to the back-to-back testing is that only the intended outputs of the systems need to match, but the inputs are different.

3.13

medical device

see IEV 871-06-06

3.14

metamorphic testing

testing where the expected results are not based on the specification but are instead extrapolated from previous actual results

[SOURCE: [ISO/IEC TR 29119-11:2020 \[3\]](#), definition 3.1.45]

3.15

pseudo-oracle

independently derived variant of the test item used to generate results, which are compared with the results of the original test item based on the same test inputs

Note 1 to entry: Pseudo-oracles are a useful alternative when traditional [test oracles \(3.19\)](#) are not available.

[SOURCE: [ISO/IEC TR 29119-11:2020 \[3\]](#), definition 3.1.59]

3.16

synthetic data

data that is artificially generated rather than produced by real-world events

3.17

test

activity in which a system or component is executed under specified conditions, the results are observed or recorded, and an evaluation is made of some aspect of the system or component

[SOURCE: [ISO/IEC/IEEE 29119-1:2022 \[1\]](#), definition 3.82]

3.18

test case

set of preconditions, inputs and expected results, developed to drive the execution of a test item to meet test objectives

Note 1 to entry: A test case is the lowest level of test implementation documentation (i.e. test cases are not made up of test cases) for the test level or test type for which it is intended.

Note 2 to entry: Test case preconditions include the required state of the test environment, data (e.g. databases) used by the test item, and the test item itself.

Note 3 to entry: Inputs are the data information and actions, where applicable, used to drive test execution.

[SOURCE: [ISO/IEC/IEEE 29119-1:2022 \[1\]](#), definition 3.85]

3.19

test oracle

source of information for determining whether a test has passed or failed

Note 1 to entry: The test oracle is often a specification used to generate expected results for individual [test cases \(3.18\)](#), but other sources may be used, such as comparing actual results with those of another similar program or system or asking a human expert.

[SOURCE: [ISO/IEC/IEEE 29119-1:2022 \[1\]](#), definition 3.115]

3.20

test specification

complete documentation of the test design, [test cases \(3.18\)](#) and test procedures for a specific test item

Note 1 to entry: A test specification can be detailed in one document, in a set of documents, or in other ways, for example in a mixture of documents and database entries.

[SOURCE: [ISO/IEC/IEEE 29119-1:2022 \[1\]](#), definition 3.125]

3.21

transparency

property of a system that appropriate information about the system is made available to relevant stakeholders

Note 1 to entry: Appropriate information for system transparency can include aspects such as features, performance, limitations, components, procedures, measures, design goals, design choices and assumptions, data sources and labelling protocols.

Note 2 to entry: Inappropriate disclosure of some aspects of a system can violate security, privacy or confidentiality requirements.

[SOURCE: [ISO/IEC 22989:2022 \[4\]](#), definition 3.5.15]

3.22 Terms specific to [Annex A](#)

3.22.1

model release

process, where an [AI model](#) is approved and taken into use under direct control by the manufacturer (see [original model \(3.22.3\)](#)) or where a [modified model \(3.22.2\)](#) is approved and taken into use under indirect control by the manufacturer

3.22.2

modified model

[AI model](#) that has changed after its [released model \(3.22.4\)](#) state

3.22.3

original model

[AI model](#) released directly by the manufacturer

3.22.4

released model

[AI model](#) that has been released either directly by the manufacturer (see [original model \(3.22.3\)](#)) or through automated control process under indirect control by the manufacturer

3.22.5

targeted release

[released model \(3.22.4\)](#) the scope of which is a subset of the scope of the [original model \(3.22.3\)](#)

4 Test planning

4.1 Test plan

The manufacturer shall establish a test plan (or plans) for conducting the activities of the testing appropriate to the purpose and test scope of the [AI component](#) under [test](#).

The goal of the test planning should be to identify and define the following:

- Test objectives