DISCLAIMER:

You may not need this template. It's aimed companies that develop devices including ML algorithms. Why do they need this template?

Due to the "blackbox" nature of many of those algorithms, their functionalities typically can't be verified in the same sense as code reviews can verify correct coding in the world of old-school software development. That's why they can only be validated by subjecting the model to rigorous testing on separate test datasets. This report is the place to document all the results (e.g. sensitivity, specificity, accuracy, ROC curve values, discussion of outlayers, edge cases, etc.). Side note: there's a disconnect between what data science and regulatory affairs understands as validation and verification. Don't get confused!

1. General Information

This document describes the specification of our machine learning (algorithm) model and its testing strategy.

Keep in mind that the results of the algorithm validation should also serve as input to and be referenced in your clinical evaluation / performance evaluation report.

Regulatory references:

ISO 13485:2016 Section	Document Section
6.2; 7.3.2	2.1
7.3.7	2.3; 3.1; 3.2
4.2.3; 5.2; 7.2.1; 7.3.3	2.4
4.1; 7.3.6	3
7.3.2; 7.3.6; 7.3.7	5

ISO 14971:2019 Section	Document Section
4.3	2.1
5.2	2.4

ISO 82304:2016 Section	Document Section
6.1	2.1

ISO 62366-1:2015 Section	Document Section
5.1; 5.2	2.4

ISO 62304:2006 Section	Document Section
5.3; 8.1.2	2.2
5.2	2.4

ISO/IEC 24028:2020 Section	Document Section
9.8.1; 10.5	3.1; 3.2
9.8.2.1	3.2
9.8.2.2	4

Relevant other documentation:

- SOP Software Development
- Software Development and Maintenance Plan
- SOUP List

2. Development Resources

For more context information, refer to the device's software development and maintenance plan.

2.1. Developer Team

Name	Role	Qualification
(...)	(...)	(...)

2.2. SOUP / Frameworks

Name	Version
PyTorch	(...)

2.3. Data

Count	Description
	<e.g. annotated heart rate dataset from a wearable>

2.4. Development Planning

Optional: remove or add to the following sub-paragraphs as appropriate.

Intended Purpose

Reference intended use; describe intended task the model should solve.

(...)

Clinical Environment

Reference clinical evaluation plan; describe clinical context in which the model will be used.

(...)

Software Architecture

Reference software architecture (diagram); describe how the model will be embedded in the overall system and how it will interact with other software components.

(...)

Quality Requirements

Reference clinical evaluation plan; describe quality criteria that shall apply to the model, for example, quantify values for sensitivity / specificity, AUC, positive/negative predictive value, accuracy, repeatability. Also, provide justification for why certain parameters may not apply and compare your approach to the current state of the art / technology.

(...)

3. Data Management

3.1. Data Acquisition

Describe how you acquired your data (sources, inclusion / exclusion criteria, possible biases, data protection measures, etc). Give a rough estimate of what size of datasets are required.

If applicable, reference relevant QMS processes.

(...)

3.2. Data Annotation

Describe how data is labeled and how you determined the ground truth. Further, you can describe: labeling requirements (quality criteria), qualification / training of staff involved, quality assurance for correct labeling. If applicable, reference relevant QMS processes.

(. . .)

3.3. Data Pre-Processing

Describe how you pre-processed data. Possible steps may include: anonymization / pseudonymization, removing data elements, converting data formats or units, handling missing values or data outliers. How did you divide data into training, validation and test data. If applicable, reference relevant QMS processes.

(. . .)

4. Algorithm Model Training and Description

Describe your algorithm model. For example: which type of model (e.g. CNN) is used? Optionally, add a graph or diagram to explain the model architecture. Describe model training; e.g. selection of features, (most important) hyperparameters, etc.

(. . .)

5. Algorithm Model Evaluation

Describe how testing is done: describe your test data set, which metrics are used for evaluation and which values are deemed acceptable.

(. . .)

6. Conclusion

Discuss your evaluation results: why do they meet predefined quality criteria, why is the model fit for purpose in accordance with the medical device's intended use? Take into account possible risks, biases and limitations.

(. . .)
