

## **SOFTWARE AS A MEDICAL DEVICE (SAMD): CLINICAL EVALUATION**

*This draft guidance, when finalized, will represent the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff or Office responsible for this guidance as listed on the title page.*

For questions about this draft document, contact Bakul Patel 301-796-5528.



**IMDRF** International Medical  
Device Regulators Forum

## PROPOSED DOCUMENT

### **International Medical Device Regulators Forum**

**Title:** Software as a Medical Device (SaMD): Clinical Evaluation

**Authoring Group:** Software as a Medical Device Working Group

**Date:** 5 August 2016

1	<b>Table of Contents</b>	
2	<b>1.0 Introduction</b> .....	<b>4</b>
3	<b>2.0 Scope</b> .....	<b>6</b>
4	<b>3.0 References</b> .....	<b>7</b>
5	<b>4.0 Definitions</b> .....	<b>8</b>
6	4.1 Clinical Validity of a SaMD .....	8
7	4.2 Scientific Validity of a SaMD .....	9
8	4.3 Clinical Performance of a SaMD.....	10
9	4.4 Analytical Validity of a SaMD .....	11
10	<b>5.0 General Principles and Context of SaMD Clinical Evaluation</b> .....	<b>12</b>
11	5.1 Clinical Evaluation Principles .....	15
12	<b>6.0 SaMD Clinical Evaluation Methods, Evidence and Appraisal</b> .....	<b>17</b>
13	6.1 What are the Evidence Goals of Clinical Evaluation?.....	17
14	6.2 Determining the Required Level of Clinical Evaluation .....	18
15	6.3 Generating Scientific Validity Evidence for a SaMD .....	19
16	6.4 Generating Analytical Validity Evidence for a SaMD .....	22
17	6.5 Generating Clinical Performance Evidence for a SaMD.....	23
18	6.6 Appraisal of Clinical Evaluation Evidence .....	24
19	<b>7.0 Level of Evidence According to SaMD Category</b> .....	<b>27</b>
20	7.1 Categories of SaMD .....	27
21	7.2 Importance of Clinical Evidence and Expectations by SaMD Category.....	27
22	7.3 Importance of Independent Review of Evidence by SaMD Category .....	29
23	7.4 Pathway for Continuous Learning Leveraging Real World Clinical Evidence.....	31
24	<b>8.0 Appendices</b> .....	<b>33</b>
25	8.1 SaMD Definition Statement .....	33
26	8.2 Clarifying SaMD Definition .....	35
27	8.3 SaMD Categorization .....	36
28	8.4 Illustrative Examples of Clinical Evaluation Concepts for SaMD.....	39
29	8.5 Summary of SaMD Clinical Evaluation recommendation .....	44
30	8.6 Glossary of Terms Interpreted for SaMD from GHTF Documents .....	45
31		

32 **Preface**

33 The document herein was produced by the International Medical Device Regulators Forum  
34 (IMDRF), a voluntary group of medical device regulators from around the world. The document  
35 has been subject to consultation throughout its development.

36 There are no restrictions on the reproduction, distribution or use of this document; however,  
37 incorporation of this document, in part or in whole, into any other document, or its translation  
38 into languages other than English, does not convey or represent an endorsement of any kind by  
39 the International Medical Device Regulators Forum.

40

## 41 **1.0 Introduction**

42 The International Medical Device Regulators Forum (IMDRF) seeks to establish a common and  
43 converged understanding of clinical evaluation and principles for demonstrating the safety,  
44 effectiveness and performance of software intended for medical purposes as defined in the  
45 [IMDRF/SaMD WG/N10](#) document *Software as a Medical Device (SaMD): Key Definitions*  
46 ([SaMD N10](#)).

47 For all medical devices, clinical evaluation, a process activity that is conducted during a  
48 product's lifecycle as part of the quality management system, is the assessment and analysis of  
49 clinical data pertaining to a medical device to verify its safety, effectiveness and performance.<sup>1</sup>  
50 The principles for clinical evaluation are the same for all medical devices and the expected rigor  
51 in current clinical guidance is intended to be technology agnostic.

52 SaMD, a type of medical device, also has significant patient and public health impact and  
53 therefore requires reasonable assurance of safety, effectiveness and performance.

54 This assurance for a SaMD is expected to be provided through a systematically planned clinical  
55 evaluation approach that generates adequate scientific evidence to create transparency, and to  
56 assure confidence in the SaMD's clinical validity for the intended purpose and indications for  
57 use, namely the claims, of the SaMD. This evaluation along with the evidence helps demonstrate  
58 that the SaMD is safe, that it performs as intended, and that the risks associated with the use of  
59 the SaMD are acceptable when weighed against the benefits to patients.

60 Global regulators expect that clinical evaluation and the evidence generated for a SaMD have the  
61 same scientific level of rigor that is commensurate with the risk and impact of the SaMD, to  
62 demonstrate assurance of safety, effectiveness and performance.

63 SaMD however is unique in that it operates in a complex highly connected-interactive socio-  
64 technical environment in which frequent changes and modifications can be implemented more  
65 quickly and efficiently. Development of SaMD is also heavily influenced by new entrants  
66 unfamiliar with medical device regulations and terminology developing a broad spectrum of  
67 applications.

68 Most SaMD's, except in limited cases, do not directly affect or have contact with a patient,  
69 instead only performs computation on data input and provides data output to a user to inform  
70 clinical management, drive clinical management, or in the diagnosis or treatment of the patient.  
71 Data input received by a SaMD typically relies on other physiological measuring medical device  
72 output or an in-vitro diagnostic device. However as healthcare decisions increasingly rely on  
73 information provided by the output of SaMD, these decisions can impact clinical outcomes and  
74 patient care.

---

<sup>1</sup> Global Harmonization Task Force (GHTF) [Clinical Evaluation](#), Page 4, May 2007.  
<http://www.imdrf.org/docs/ghrf/final/sg5/technical-docs/ghrf-sg5-n2r8-2007-clinical-evaluation-070501.pdf>

75 Based on the significant impact SaMD has on clinical outcomes and patient care, a SaMD  
76 manufacturer is expected to gather, analyze, and evaluate data, and develop evidence to  
77 demonstrate the assurance of safety, effectiveness and performance of the SaMD. This  
78 evaluation should focus on how well the information provided by the SaMD meets the clinical  
79 needs within the intended healthcare situation and condition that includes consideration for the  
80 target population, characteristics of the disease or condition, and type of user. This document  
81 discusses addressing these clinical needs by demonstrating the analytical validity (the SaMD's  
82 output is accurate for a given input), and where appropriate, the scientific validity (the SaMD's  
83 output is associated to the intended clinical condition/physiological state), and clinical  
84 performance (the SaMD's output yields a clinically meaningful association to the target use of  
85 the SaMD) of the SaMD.

86 In addition to these general clinical evaluation expectations, this guidance considers the  
87 uniqueness of indirect contact between patients and SaMD and presents the principles of clinical  
88 evaluation with recommendations to address this uniqueness. Additionally, this document  
89 highlights the uniqueness of SaMD that can leverage the connected-interactive socio-technical  
90 environment to continuously learn from real world use information. SaMD manufacturers can  
91 use this real world information to support the assurance of safety, effectiveness and performance,  
92 in a continuous and agile clinical evidence gathering paradigm. This paradigm shifts the focus  
93 towards observed real world performance as part of post-market monitoring.



*Clinical evaluation is the assessment and analysis of clinical data pertaining to a medical device in order to verify the safety, effectiveness and performance of the device. Clinical evaluation is an ongoing process conducted during the lifecycle of a medical device.*

94 This document primarily references previous Global Harmonization Task Force (GHTF<sup>2</sup>) and  
95 IMDRF guidance documents to provide a common understanding and application of  
96 terminology, concepts and principles for performing a clinical evaluation to demonstrate the  
97 performance of a SaMD.

98 This application of clinical evaluation principles and concepts for a SaMD also relies on the  
99 principles and processes described in [IMDRF IMDRF/SaMD WG/N23FINAL:2015](#) Application  
100 of Quality Management Systems (QMS) (*SaMD N23*). Specifically [SaMD N23](#) describes how  
101 clinical evaluation is also a process within the lifecycle activities, and the larger quality  
102 management systems framework that includes organizational support, lifecycle support processes  
103 and realization software development lifecycle processes.

---

<sup>2</sup> GHTF was a voluntary group of representatives from national medical device regulatory authorities and industry representatives. GHTF was disbanded in 2012 and its mission has been taken over by the IMDRF.

104 As with other medical devices, the level of documented clinical evidence expected by a regulator  
105 will depend on regulatory laws in their individual jurisdictions where the SaMD is intended to be  
106 made available. This document does not opine on the individual jurisdiction's requirement;  
107 instead this document provides guidance on the relative importance and expectations, based on  
108 the impact to health, for conducting clinical evaluation and documented evidence for the  
109 different categories of SaMD as described in [IMDRF IMDRF/SaMD WG/N12FINAL:2014](#)  
110 ([SaMD N12](#)).

111 This is a companion document to *SaMD N10*, *N12* and *N23* documents, further enabling  
112 convergence in vocabulary, approach, and a common thinking for regulators and industry. It  
113 should also be noted that this document does not provide guidance on the adequacy of meeting  
114 regulatory requirements or "essential principles" that are the basis of GHTF classifications.  
115 Rather this guidance provides the relative importance of required clinical performance for the  
116 different categories of SaMD as categorized in the [SaMD N12](#) document.

## 117 **2.0 Scope**

118 The objective of this document is to provide guidance on clinical evaluation by describing:

- 119 • Relevant clinical evaluation methods and processes which can be appropriately used for  
120 SaMD to generate clinical evidence;
- 121 • The necessary level of clinical evidence for different categories of SaMD; and
- 122 • SaMD categories where independent review is important or not important.

123 The principles discussed are intended to assist SaMD manufacturers and regulators. The  
124 principles are based on a common goal to provide confidence to the users of SaMD (patients,  
125 providers, consumers, clinical investigators) who rely on the output of SaMD for patient care.

126 The description of appropriate clinical evaluation methods and processes for SaMD, and  
127 recommendations for how much evidence (or degree of certainty of the evidence), and  
128 independent oversight is appropriate for SaMD, is not meant to replace or conflict with pre-  
129 market or post-market regulatory requirements related to the regulatory classification of SaMD  
130 in different jurisdictions. Similarly, the information is not meant to replace, or conflict with,  
131 technical or international standards.

132 In achieving the above objectives, this document relies upon and does not repeat the concepts  
133 and principles found in [SaMD N12](#) (risk categorization of SaMD), and [SaMD N23](#) (application  
134 of quality management for SaMD), but is a continuum to those documents, and this document  
135 should be used in conjunction with those.

136 The categories of SaMD are limited to the definition in [SaMD N10](#) and the categories of intended  
137 use described in [SaMD N12](#) where the information provided by SaMD is intended to inform  
138 clinical management, drive clinical management, or diagnose or treat a disease or condition in  
139 non-serious, serious or critical healthcare situations or conditions.

140

141

Not SaMD	SaMD Categories									Not SaMD
Retrieves Information	I			II			III		IV	Software used in Closed Loop Interventions
Organizes Data	Informs Non-Serious	Informs Serious	Drives Serious	Informs Critical	Drives Serious	Treats/ Diagnoses Non Serious	Drives Critical	Treats/ Diagnoses Serious	Treats/ Diagnoses Critical	
Optimizes Processes										

Figure 1- What is / is not a SaMD

142  
 143 Note: Refer to Sections 8.2 and 8.3 for more information and examples related to what is a SaMD  
 144 and what is not a SaMD.

145 This document specifically does not include in its scope or address other types of software used  
 146 in health care for retrieving information from devices or systems, organizing the collected data,  
 147 or optimizing healthcare workflow by automating healthcare provider’s care protocols. The  
 148 scope of SaMD also does not include software that is embedded in a physical medical device or  
 149 software that is used to provide closed loop intervention (see [Section 9.1 Clarifying SaMD](#)  
 150 [Definition](#) for more information and examples).

151 The guidance provided in this document specifically does not address the regulatory  
 152 classification of SaMD and does not address whether a premarket clearance is required for a  
 153 specific SaMD.

154 This guidance also does not address issues that are generic to all medical devices or specific to a  
 155 country or jurisdiction such as the following:

- 156 • Off-label use or foreseeable misuse;
- 157 • Device classification of specific SaMD;
- 158 • Whether a pre-market approval or certification is required for specific SaMD.

159 **3.0 References**

160 **IMDRF Documents:**

- 161 SaMD N10 [Software as a Medical Device \(SaMD\): Key Definitions](#)
- 162 SaMD N12 [Software as a Medical Device \(SaMD\): Possible Framework for Risk](#)  
[Categorization and Corresponding Considerations](#)
- 163 SaMD N23 [Software as a Medical Device \(SaMD\): Application of Quality](#)  
[Management System](#)

166 **GHTF Documents:**

- 167 GHTF SG5 /N6 [Clinical Evidence for IVD medical devices – Key Definitions and](#)  
[Concepts](#)
- 169 GHTF SG5 /N7 [Clinical Evidence for IVD medical devices - Scientific Validity](#)  
[Determination and Performance Evaluation](#)
- 171 GHTF SG5 /N8 [Clinical Evidence for IVD Medical Devices - Clinical Performance](#)  
[Studies for In Vitro Diagnostic Medical Devices](#)



---

173	GHTF SG5 /N3	<a href="#">Clinical Investigations</a>
174	GHTF SG5 /N2	<a href="#">Clinical Evaluation</a>
175	GHTF SG5 /N1	<a href="#">Clinical Evidence – Key Definitions and Concepts</a>
176	GHTF SG5 /N4	<a href="#">Post-Market Clinical Follow-up Studies</a>
177	GHTF SG1 /N68	<a href="#">Essential Principles of Safety and Performance of Medical Devices</a>

178

**179 International Standards:**

180	ISO 14155-1:2011	<a href="#">Clinical investigation of medical devices for human subjects -- Good clinical practice</a>
181		
182	ISO 14971:2007	<a href="#">Application of risk management to medical devices</a>
183	IEC 80002-1:2009	<a href="#">Medical device software -- Part 1: Guidance on the application of ISO 14971 to medical device software</a>
184		

**185 4.0 Definitions**

186 This document does not introduce any new definitions but rather relies on the following:

- 187 • Definition of SaMD as identified in [SaMD N10](#).

188 *Software as a Medical Device (SaMD)*

189 *The term “Software as a Medical Device” (SaMD) is defined as software intended to be*  
190 *used for one or more medical purposes that perform these purposes without being part*  
191 *of a hardware medical device.*

192 *NOTES:*

- 193 ○ *SaMD is a medical device and includes in-vitro diagnostic (IVD) medical device.*  
194 ○ *SaMD is capable of running on general purpose (non-medical purpose) computing*  
195 *platforms*  
196 ○ *“without being part of” means software not necessary for a hardware medical*  
197 *device to achieve its intended medical purpose;*  
198 ○ *Software does not meet the definition of SaMD if its intended purpose is to drive a*  
199 *hardware medical device.*  
200 ○ *SaMD may be used in combination (e.g., as a module) with other products including*  
201 *medical devices;*  
202 ○ *SaMD may be interfaced with other medical devices, including hardware medical*  
203 *devices and other SaMD software, as well as general purpose software*  
204 ○ *Mobile apps that meet the definition above are considered SaMD.*

- 205 • Definition of Clinical Evaluation and associated terms and vocabulary as identified by the  
206 Global Harmonization Task Force (GHTF) and interpreted for a SaMD not included in  
207 Section 4.0 Definitions below can be found in Appendix A of this document.

**208 4.1 Clinical Validity of a SaMD**

209 For purposes of this guidance, the term clinical validity is used to refer to the combination of:

- 210 a) The association of the output of a SaMD to a clinical condition/physiological state  
211 (scientific validity); together with  
212 b) The ability of a SaMD to yield a clinically meaningful output associated to the target  
213 use of SaMD output in the health care situation or condition identified in the SaMD  
214 definition statement (clinical performance).

215 Depending on the type of SaMD, clinical validity can be expressed as follows:

- 216 • For SaMD that is intended to treat a disease or condition, clinical validity is the evidence  
217 of effectiveness of the SaMD output to the treatment or prevention.
- 218 • For non-diagnostic SaMD, clinical validity is the evidence of scientific validity that  
219 shows the usefulness of the SaMD output in clinical care.
- 220 • For diagnostic SaMD, clinical validity is the evidence of scientific validity in addition to  
221 the clinical performance evidence of the SaMD.

## 222 4.2 Scientific Validity of a SaMD

223 Scientific validity is the association of the SaMD output to a clinical condition/physiological  
224 state.

225 Scientific validity is often identified from academic research, and is often supported by studies  
226 evaluating the inputs along with the algorithms for an association of the SaMD's output to a  
227 clinical condition/physiological state. *Example: Hemoglobin concentration is associated with*  
228 *anemia (clinical condition). Congestive heart failure, Hypertension, Age, Diabetes, prior Stroke*  
229 *(CHADS-2) score is associated with predicting the risk of stroke in patients with non-valvular*  
230 *atrial fibrillation.*

231 Scientific validity establishes how well the output of the SaMD accurately correlates to the  
232 intended clinical health care situation or condition of the intended use of the SaMD. The  
233 evidence demonstrates objectively the clinical association of the SaMD's use of inputs,  
234 algorithm and outputs as compared to a recognized reference standard (i.e., gold standard), to  
235 another SaMD or medical device, to a well-documented method, to the current clinical practice  
236 or standard of care, or as compared to a composite reference standard. When comparing to other  
237 devices, including other SaMD's, the original reference standard used by the other device to  
238 determine the scientific validity of the intended clinical condition is typically used rather than the  
239 device itself.

240 Scientific validity also determines if the association of the SaMD's intended use to a clinical  
241 condition/physiological state is well-known (i.e., known clinically acceptable analytical validity  
242 standards, and where the analytical validity assessment has determined that the SaMD meets  
243 those standards), based on available review of information such as peer reviewed literature,  
244 textbooks, historical data and experience based evidence, academic research, or is supported by  
245 previous studies.

246 At the conclusion of scientific validity appraisal, a SaMD can generally be segregated in one of  
247 the following categories:

- 248 a) **Well-known association:** These SaMD's have output with a well-known association to  
249 identified clinical guidelines, clinical studies in peer reviewed journals, consensus for the use  
250 of the SaMD, international reference materials or other similar sources. *Example:*

251 *Computation of the Acute Physiology and Chronic Health Evaluation II (APACHE II) score*  
252 *is a well-known association to stroke risk.*

253 b) **Novel association:** These SaMD's involve, new inputs, algorithms or outputs, new intended  
254 target population, or a new intended use, and they are not well-known. *Example(s): use of*  
255 *non-standard input such as gait, blood pressure or other physiological and environmental*  
256 *signals using novel algorithms to detect early onset of a deterioration of health or diagnosis*  
257 *of a disease. .*

#### 258 4.3 Clinical Performance of a SaMD

259 The clinical performance of a SaMD is the ability of a SaMD to yield a clinically meaningful  
260 output associated to the target use of SaMD output in the health care situation or condition  
261 identified in the SaMD definition statement (disease type, target user, and intended population).  
262 Clinically meaningful means the positive impact of a SaMD on the health of an individual, to be  
263 specified as meaningful, measurable, patient-relevant clinical outcome(s), including outcome(s)  
264 related to diagnosis or a positive impact on patient management of public health.

265 Clinical performance is evaluated and determined by the manufacturer during the development  
266 of a SaMD before it is distributed for use (pre-market) or after distribution while the SaMD is in  
267 use (post-market).

268 Clinical performance of a SaMD can also be viewed as the relationship between the verification  
269 and validation results of the SaMD algorithm and the clinical conditions of patients. This  
270 performance can also be determined using real world data, where the data is useful in identifying  
271 less common use situations.

272 The clinical performance of a SaMD may be characterized by demonstrating:

- 273 • Sensitivity - ability of the SaMD to correctly identify across a range of available  
274 measurements patients with the intended clinical disease or condition;
- 275 • Specificity - ability of a SaMD to correctly identify across a range of available  
276 measurements patients that do not have the intended disease or condition;
- 277 • ROC curve - a graphical plot that shows the tradeoff between sensitivity and specificity  
278 as the decision threshold that separates SaMD's negatives and positives is varied;
- 279 • Positive predictive value – which indicates the likelihood of the patient having a disease  
280 or condition given that the SaMD's output is positive;
- 281 • Negative predictive value – which indicates the likelihood of the patient NOT having a  
282 disease or condition given that the SaMD's output is negative;
- 283 • Likelihood ratio - the likelihood that a given result would be expected in a patient with  
284 the target condition compared to the likelihood that the same result would be expected in  
285 an individual without that condition; and
- 286 • Cut-off thresholds, indices or scales – should be meaningful for the intended use of the  
287 SaMD and established prior to validation.

288 **NOTE:** The sensitivity and specificity depend on the choice of a cut-off value (e.g., to separate  
289 negative from positive values).

290 **NOTE:** Predictive value depends on the prevalence of the disease or condition in the population  
291 of interest.

---

#### 292 4.4 Analytical Validity of a SaMD

293 The analytical validity of a SaMD is the ability of a SaMD to accurately and reliably generate the  
294 intended output, from the input data, i.e., analytical validity measures the SaMD's ability to  
295 correctly and reliably process input data and generate output data with accuracy, and  
296 repeatability and reproducibility, i.e., precision. Analytical validity may also include measures  
297 for analytical sensitivity (e.g., limit of detection), and linearity or behavior of output across the  
298 range of input data that is allowed by the SaMD.

299 Analytical validity is generally evaluated and determined by the manufacturer during the  
300 verification and validation phase of the software development lifecycle using a QMS. Analytical  
301 validity is always expected for a SaMD.

302 Analytical validity confirms and provides objective evidence that (a) the software meets its  
303 specification, in other words, "is the software being built right?", and (b) software specifications  
304 conform to user needs and intended uses, and that the particular requirements implemented  
305 through software can be consistently fulfilled, in other words, "is the right software being built?"

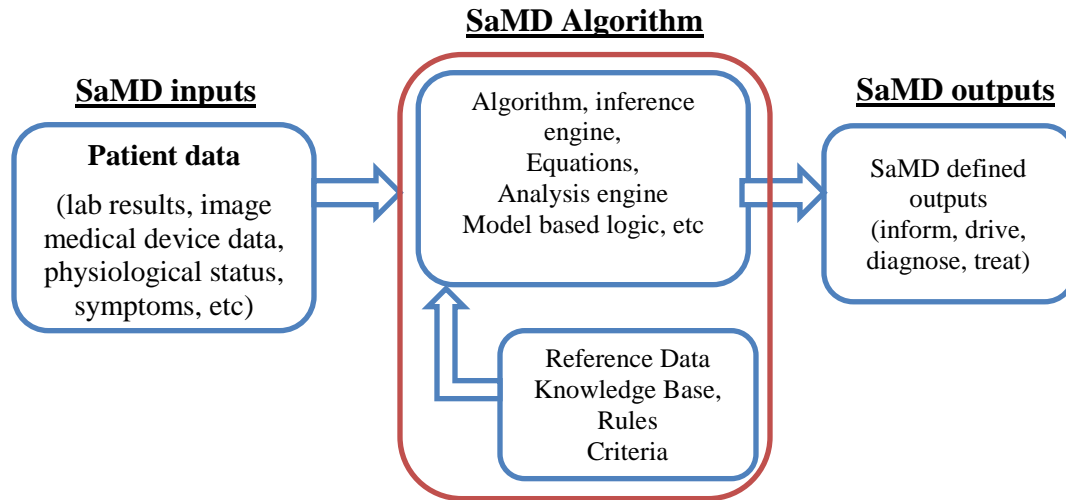
306 The analytical validity of a SaMD will include measures to demonstrate the following:

- 307 • Accuracy - degree of closeness of measurements of a quantity to that quantity's true  
308 value. When the output of the SaMD and true value are binary, accuracy is the proportion  
309 of true results (both true positives and true negatives) among the total number of output  
310 values examined;
- 311 • Precision - related to reproducibility and repeatability, is the degree to which repeated  
312 measurements under unchanged conditions show the same results;
- 313 • Limit of detection - ability of the SaMD to discern between information-bearing patterns  
314 of a clinical condition and random patterns that distract from the information;
- 315 • Linearity or associated transfer function - the behavior of the output across the range of  
316 input data that is allowed by the SaMD; and
- 317 • Analytical sensitivity - degree to which the algorithm's output is affected by the input  
318 data (e.g., parameters affecting input data may include perturbation, image resolution,  
319 illuminations, data spatial distribution, data amount, etc.).

320

321 **5.0 General Principles and Context of SaMD Clinical Evaluation**

322 At the highest and simplest level of abstraction a SaMD can be described as a software that  
 323 utilizes an algorithm (logic, set of rules, or a model) that operates on data input (digitized  
 324 content) to produce an output that is information intended for medical purposes as defined by the  
 325 SaMD manufacturer as represented in Figure 2 below.

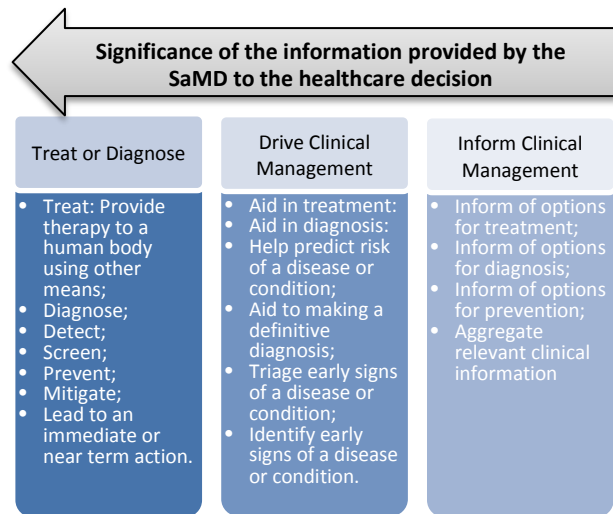


326

327 **Figure 2: High Level SaMD Components**

328 The risks and benefits posed by a SaMD are largely related to the risk of the output of the SaMD  
 329 if not accurate (or correct) which in turn  
 330 impacts the clinical management of a  
 331 patient; rather than the risk from direct  
 332 contact between the SaMD and the patient.  
 333 As covered in SaMD Risk Framework (())  
 334 many aspects affect the importance of the  
 335 output information from SaMD. Generally  
 336 these aspects can be grouped into the  
 337 following two major factors that provide  
 338 adequate description of the intended use of  
 339 SaMD:

- 340 A. Significance of the information
- 341 provided by the SaMD to the
- 342 healthcare decision, and
- 343 B. State of the healthcare situation or
- 344 condition.



345 **Figure 3 – SaMD N12 components of "significance" of SaMD output (See Section 8.1 of this document)**

345

← State of the healthcare situation or condition

	Critical	Serious	Non-Serious
<i>Disease Type /Patient Condition</i>	<ul style="list-style-type: none"> <li>• Life-threatening;</li> <li>• Incurable</li> <li>• Fragile</li> </ul>	<ul style="list-style-type: none"> <li>• Moderate in progression</li> <li>• Often curable;</li> <li>• Not fragile;</li> </ul>	<ul style="list-style-type: none"> <li>• Slow with predictable progression of disease state</li> <li>• Minor chronic illnesses or states</li> <li>• May not be curable;</li> <li>• Individuals who may not always be patients</li> <li>• Can be managed effectively</li> </ul>
<i>Intervention Type</i>	<ul style="list-style-type: none"> <li>• Requires major therapeutic interventions;</li> <li>• Sometimes time critical</li> <li>• Vital to: avoiding death; serious deterioration of health; mitigating public health situations or conditions</li> </ul>	<ul style="list-style-type: none"> <li>• Does not require major therapeutic interventions</li> <li>• Not expected to be time critical</li> <li>• Vital to avoiding unnecessary interventions</li> </ul>	<ul style="list-style-type: none"> <li>• Slow with predictable progression of disease state</li> <li>• Minor chronic illnesses or states;</li> <li>• May not be curable;</li> <li>• Individuals who may not always be patients</li> <li>• Can be managed effectively</li> </ul>
<i>User Type</i>	<ul style="list-style-type: none"> <li>• Specialized trained users</li> </ul>	<ul style="list-style-type: none"> <li>• Either specialized trained users or lay users.</li> </ul>	<ul style="list-style-type: none"> <li>• Either specialized trained users or lay users</li> </ul>

347

365

When these factors are included in the manufacturer’s description of intended use, they can be used to categorize SaMD. [SaMD N12 Section 6.0](#) provides a structured approach for a SaMD definition statement to describe the intended use. [SaMD N12 Section 7.0](#) provides a method for categorizing SaMD based on the major factors identified in the definition statement. (See section 8.3 for the SaMD categorization)

In limited cases -- where SaMD may have the functionality to accept user inputs or to “treat” using general purpose computer peripherals to impart sound, light, pictures on a display or in some cases low energy vibrations -- such SaMD can be considered to provide therapy to patients (e.g., SaMD used for cognitive behavioral therapy).

These categories include functionality that has an increasing significance of the output to the patient care.

368 Illustrative examples of SaMD along this spectrum include:

- 369 • A SaMD that performs analysis of cerebrospinal fluid spectroscopy data to diagnose  
370 tuberculosis meningitis or viral meningitis in children. Such SaMD is used to diagnose a  
371 disease in a fragile population with possible broader public health impact that may be life  
372 threatening, may require major therapeutic intervention, and may be time sensitive  
373 (*SaMD N12 Category IV.i*).
- 374 • SaMD that is intended as a radiation treatment planning system as an aid in treatment in a  
375 critical condition that may be life threatening and requires major therapeutic intervention  
376 (*SaMD N12 Category III.ii*).
- 377 • SaMD that uses data from individuals for predicting risk score for developing stroke or  
378 heart disease for creating prevention or interventional strategies (*SaMD N12 Category  
379 II.iii*).
- 380 • SaMD that analyzes images, movement of the eye or other information to guide next  
381 diagnostic action of astigmatism. Such SaMD provides aggregation of data to provide  
382 clinical information that will not trigger an immediate or near term action for the treatment  
383 of a patient condition that even if not curable can be managed effectively and whose  
384 interventions are normally noninvasive in nature (*SaMD N12 Category I.i*).

385 Other aspects that affect the safety, effectiveness and performance of a SaMD include  
386 considerations for:

- 387 • Socio-technical environment consideration ([SaMD N12](#) Section 9.1) when identifying  
388 effects/implications and appropriate measures for safety, effectiveness and performance  
389 of SaMD throughout the product's design, development and installation including:
  - 390 ○ Usability of the application - How integrating SaMD within real-world clinical  
391 workflows.
  - 392 ○ Transparency of the inputs, outputs and methods to the user.
- 393 • Technology and system environment consideration ([SaMD N12](#) Section 9.2).
- 394 • Information security with respect to safety consideration ([SaMD N12](#) Section 9.3).

395 These other aspects influence the identification of considerations that are unique to a specific  
396 approach/method used by the manufacturer of a particular category of SaMD. For example, the  
397 type of a platform, that is constantly changing, used in the implementation of SaMD may create  
398 considerations that are unique to that implementation. These considerations can also vary by the  
399 capabilities of the manufacturer or by the process rigor used to implement the SaMD. This rigor  
400 as outlined in N23 expects that all manufacturers of SaMD follow adequate QMS that include  
401 risk management processes to manage technological, use environment and clinical risks.

- 402 • The governance structure ([SaMD N23](#) Section 6.0) should provide support for creating  
403 and establishing appropriate processes that are important for maintaining the quality  
404 objectives and policies<sup>3</sup>;
- 405 • The elements of SaMD lifecycle support processes ([SaMD N23](#) Section 7.0) that are  
406 common processes and activities that should be considered throughout the SaMD  
407 lifecycle regardless of specific software product development approach or method used  
408 by the organization. These processes -- product planning; risk management: a patient  
409 safety focused approach; document and record control; configuration management and  
410 control; measurement, analysis and improvement of processes and product; managing  
411 outsourced processes and products – that should be applied throughout the SaMD  
412 realization and use processes; and
- 413 • Aspects of realization and use processes ([SaMD N23](#) Section 8.0) commonly found in  
414 software engineering lifecycle approaches (process, activities, tasks, etc.) that are  
415 important for an effective SaMD QMS include: requirements management, design,  
416 development, verification and validation, deployment, maintenance, decommissioning  
417 (retirement or end-of-life activity).

418 QMS rigor when applied correctly is expected to have adequate rigor in generating evidence  
419 towards:

- 420 • Managing uniqueness of short development cycle for SaMD development and changes  
421 ([SaMD N23](#) Section 8.6).
- 422 • Control over distribution channels ([SaMD N23](#) Section 8.5).

---

<sup>3</sup> These processes, policies and objectives should be tailored for the needs, type, size and nature of an organization.

- 423 • Controlling design/specification changes, versioning, monitoring installed base,  
424 managing recalls, remote updates ([SaMD N23](#) Section 8.5).
- 425 • Quality – Usability (including user interface), conformance to specifications, “fitness for  
426 use”, and reasonably free from the possible serious effects of defects with a plan in place  
427 to detect and correct the defects to ensure the SaMD continues to meet the intended  
428 safety, effectiveness and performance.
- 429 • Objectively verified and validated to show conformance to customer requirements.
- 430 • Managed quality while in use through timely maintenance and continuous improvement.

## 431 5.1 Clinical Evaluation Principles

432 Like other high-quality products, a SaMD manufacturer implements on-going lifecycle processes  
433 to thoroughly evaluate the product’s performance in its intended market. Prior to product launch,  
434 the manufacturer continues to collect evidence of the product’s accuracy, specificity, reliability,  
435 limitations, and scope of use in the intended use environment with the intended user. Once the  
436 product is on the market the manufacturer continues to gather evidence to further understand the  
437 customer’s needs in a real world environment and to ensure the product is meeting those needs.  
438 This real world information allows the manufacturer to identify and correct any problems and to  
439 enhance the product by expanding functionality to stay competitive or meet user demands.

440 Lifecycle activities, including clinical evaluation, should follow appropriate planning processes  
441 as part of an organization’s lifecycle activities and processes. This means clinical evaluation,  
442 similar to other SaMD lifecycle activity and process, also needs to be planned prior to  
443 conducting the evaluation. Risk assessment done as part of the SaMD’s lifecycle activities and  
444 processes should also be considered when conducting clinical evaluation. Risk, including the  
445 impact of hazards and hazardous situations identified while conducting clinical evaluation should  
446 be incorporated into the overall risk management processes of SaMD. The following are  
447 examples of considerations for risk management that may impact clinical evaluation:

- 448 • Level of clinical evidence available and the confidence of the evidence;
- 449 • Complexity of the clinical model used to derive the output information;
- 450 • Known specificity of the output information;
- 451 • Maturity of clinical basis of the software and confidence in the output;
- 452 • Benefit of the output information vs. current standard of care;
- 453 • Feasibility ([SaMD N23](#) Section 7.1);
- 454 • User and patient needs intended use ([SaMD N23](#) Section 8.3); and
- 455 • Clinical evidence that product meets clinical end user expectations ([SaMD N23](#) Section  
456 8.4).

457 SaMD clinical evaluation includes the gathering and assessment of scientific validity, analytical  
458 validity and clinical (real-world, obtained from patients) performance of a SaMD. A combination  
459 of the results of these activities generates clinical evaluation evidence for a SaMD.

460 The extent of clinical evaluation evidence necessary for a SaMD will depend on parameters  
461 including but not necessarily limited to the underlying algorithm, the transparency of the  
462 algorithm along with the ability for a user to detect erroneous output, the degree of variability of



463 the subject population and disease state (intended use target population), and the intended user(s)  
464 of the SaMD. Clinical evaluation of SaMD is expected to be iterative and continuous.

465 While not intended to impose unnecessary burden, clinical evidence should support the intended  
466 use of the SaMD as stated by the manufacturer while addressing the relative risks to the patient  
467 associated with the use of the SaMD. The intended use for a SaMD defines the medical purpose  
468 and determines the type and depth of the clinical evaluation. This statement of intention is the  
469 most important starting point for considering the level of evidence necessary and in the choices  
470 made to perform appropriate clinical evaluation.

471 For purposes of this document, performing clinical evaluation and generating data for SaMD  
472 assumes the following prerequisites:

- 473 • Clinical evaluation scope is dependent on “intended use” as defined by the manufacturer  
474 of SaMD.
  - 475 ○ The intended use of the SaMD is dependent on the product claims. The product  
476 claims, along with the SaMD definition statement determines the level of clinical  
477 evidence needed. Performance, functionality, and features as defined by the  
478 manufacturer are expected to be consistent with the claims.
  - 479 ○ While the SaMD is on the market, claims should reflect the actual performance  
480 and functionality of the SaMD (real world performance.)

481

---

## 482 **6.0 SaMD Clinical Evaluation Methods, Evidence and Appraisal**

483 Clinical evaluation is a systematic and planned process to continuously generate, collect,  
484 analyze, and assess the clinical data pertaining to a SaMD in order to verify the scientific  
485 validity, and the analytical validity and clinical performance of the SaMD when used as intended  
486 by the manufacturer. The level and extent of clinical evaluation necessary is determined by the  
487 role of the SaMD for the target clinical condition. The quality and breadth of the clinical  
488 evaluation assures that the output of the SaMD is scientifically valid and can be used reliably and  
489 predictably.

490 While a prospective (e.g., randomized controlled) trial may satisfy the requirements for real-  
491 world performance, prospective trials may not be required to generate patient data. The term  
492 ‘clinical evaluation’ should not be understood to be limited to conducting a prospective  
493 randomized clinical trial.

494 This section explains the goal of clinical evaluation in generating evidence, what techniques are  
495 available for a SaMD manufacturer to generate that evidence and when such evaluation is  
496 conducted in the product lifecycle.

### 497 **6.1 What are the Evidence Goals of Clinical Evaluation?**

498 The outcome of the clinical evaluation process of a SaMD is essential to the SaMD’s value for  
499 the user and ultimately patients. The clinical evaluation evidence of a SaMD, as expressed in the  
500 intended use by the manufacturer, is generated from and validated by performing clinical  
501 evaluation and demonstrating the following:

- 502 • Scientific validity – showing with evidence on the association of the SaMD output to a  
503 clinical condition/physiological state;
- 504 • Analytical validity – showing with evidence the technical performance related to  
505 accuracy, reliability, repeatability and reproducibility; and if necessary
- 506 • Clinical performance – typically for diagnostic SaMD (see box below), showing evidence  
507 of the ability of a SaMD to yield a clinically meaningful output associated to the target  
508 use of SaMD output in the health care situation or condition.

509 Analytical validity addresses how well the device measures what it claims to measure whereas  
510 clinical performance addresses how useful that measurement is.

511 For most SaMD the goal of clinical evaluation is to establish clinical validity and to create  
512 evidence with evaluation methods that use of patient data to understand the analytical validity  
513 and clinical performance. In most cases since SaMD’s output has an influence on a user’s  
514 decision, clinical evaluations are typically focused towards the user’s ability to use the output as  
515 intended by the manufacturer. In certain instances when SaMD is intended to treat a healthcare  
516 situation or condition, clinical evaluation is conducted using patients or data that is representative  
517 or related to the patient’s situation or condition to demonstrate effectiveness of the treatment.  
518 For example, a SaMD that is intended to provide sound therapy to treat, mitigate or reduce  
519 effects of tinnitus for which minor therapeutic intervention is useful would require that the  
520 manufacturer provide analytical validity that assures that the treatment output is in accordance  
521 with all appropriate performance specifications and limitations. The manufacturer would also

522 demonstrate that there is a well-known scientific validity that associates specified sounds with an  
523 intended treatment.

524 Generally, SaMD that is not intended for treating a situation or condition can be grouped as  
525 follows:

- 526 • **Diagnostic SaMD:** These SaMD typically differentiate patients or their physiological  
527 conditions and are intended to drive clinical management and / or diagnose. Such SaMD  
528 are typically intended to identify early signs, triage, predict risk, screen, detect or  
529 diagnose a healthcare situation or condition.
- 530 • **Non-diagnostic SaMD:** These SaMD have generic functionality that can be used across  
531 various health care situations or conditions. Such SaMD typically provide data to help aid  
532 in diagnosis, aid in treatment, inform of options. Examples of such SaMD include  
533 calculators (radiation treatment planning SaMD), search and match, filter, user defined  
534 rules based matching, processing a signal (e.g., spectral analysis of a sound signal), a  
535 memory test that gives a score but no interpretation, etc.

## 536 6.2 Determining the Required Level of Clinical Evaluation

537 Clinical evaluation is an ongoing process throughout the lifecycle of a SaMD. It is based on data  
538 collected during the pre- or post-market of the product lifecycle for the SaMD intended use.

539 During the development phase of the SaMD lifecycle, clinical evaluation allows the  
540 manufacturer to objectively assess and demonstrate that the SaMD achieves its intended purpose  
541 during normal conditions of use and the known and foreseeable risks associated with the SaMD  
542 are minimized. The residual risks are acceptable when weighed against the benefits of the SaMD  
543 based on its intended use, and that any safety, effectiveness and performance claims made about  
544 the SaMD are supported by suitable evidence. Clinical evaluation also provides opportunities to  
545 assess the SaMD design characteristics, algorithm, and technological features to optimize its  
546 clinical effectiveness while minimizing any potential risks.

547 Information related to clinical evidence should be monitored routinely by the manufacturer and  
548 user once the SaMD is available on the market. The manufacturer should plan for the continuous  
549 discovery of clinical data related to the safety, effectiveness and performance of the SaMD  
550 through appropriate post-market programs (e.g., post-market surveillance, adverse event reports,  
551 scientific publications, etc.) as part of their QMS to ensure the SaMD continues to meet the  
552 intended safety, effectiveness and performance.

553 The following provide an overview of steps for generating clinical evidence.

- 554 1. Determine if scientific validity of the SaMD is already well-known with clinically  
555 accepted analytical validity standards, and where the analytical validity assessment  
556 has determined that the SaMD meets those standards:
  - 557 • If Yes: document evidence as outlined in Section 7.2;
  - 558 • If No: generate scientific validity evidence as outlined in Section 6.3
- 559 2. Perform analytical validity: As part of SaMD verification and validation activities  
560 generate analytical evidence as highlighted in Section 6.40.
- 561 3. Establish the need for clinical performance:

- 562
- 563
- 564
- 565
- 566
- 567
- 568
- 569
- 570
- 571
- 572
- For Diagnostic SaMD and has a higher risk profile (refer to SaMD expectations in section 7.2 to determine the need) – conduct clinical performance evaluation as outlined in Section 7.2.
4. If clinical performance evidence is necessary, and scientific validity is not well-known the following questions should be considered when planning the clinical performance evaluation for a SaMD:
- Is patient data available to conduct performance evaluation or is new patient data required to support the intended claim?
  - If new patient data is necessary to support the claim, what type of clinical performance evidence is necessary to pursue?
  - Refer to Section 6.3.2 below for approaches and considerations.

### 573 **6.3 Generating Scientific Validity Evidence for a SaMD**

574 Generating scientific validity evidence for a SaMD is not necessary where association of a  
575 SaMD's output to a clinical condition/physiological state is already well-known, based on  
576 available information. An example of a well-known association is the Congestive heart failure,  
577 Hypertension (CHADS-2) score used for risk stratification of ischemic stroke in patients with  
578 non-valvular atrial fibrillation.

579 Scientific validity evidence should be derived from a critical appraisal of its merits and  
580 limitations and appraised to determine each piece of information on its relevance and quality for  
581 establishing the association between the output and algorithm of the SaMD and the clinical  
582 condition/physiological state.<sup>4</sup> Scientific validity evidence for a SaMD can be generated from  
583 following methods:

- 584
- Conducting literature search:
    - Review of information found in peer reviewed articles, regulatory guidance documents, conference proceedings, case reports, etc.; Literature sources used to identify data may include: Scientific databases; specialized databases; systematic review databases; clinical trial registers; or reference tests.
    - Review of expert opinions: this information might be found in sources that include textbooks, clinical guidance documents, and position statements from academic and professional organizations;
    - Results from proof of concept studies: these studies are usually smaller scale scientific studies to identify the fundamental association of the algorithm with the clinical condition/physiological state;
    - Results from previously conducted clinical studies that provide association of a signal or output of an algorithm with a healthcare situation or condition or a physiological state.
- 585
- 586
- 587
- 588
- 589
- 590
- 591
- 592
- 593
- 594
- 595
- 596
- 597

---

<sup>4</sup> See GHTF SG5 /N7:2012 [Scientific Validity Determination and Performance Evaluation](#) Section 6.0 and 7.2.2, 7.2.3 Scientific Validity Determination for additional details related to potential sources for the identification of scientific validity information and the appraisal and analysis of scientific validity information

- 598
- 599
- 600
- 601
- 602
- 603
- 604
- 605
- 606
- 607
- 608
- 609
- 610
- 611
- 612
- 613
- 614
- 615
- 616
- 617
- 618
- Identifying scientific validity from manufacturers experience data<sup>5</sup>
    - Customer feedback including complaints, adverse events, and other data that can be systematically and scientifically provides an association of the intended SaMD output with a healthcare situation or condition or a physiological state
    - Real world data generated outside of clinical performance studies provides real world experience obtained in larger, heterogeneous and more complex intended use scenarios. The data are most useful in identifying less common but potentially serious SaMD related adverse events. The source of this additional data may include:
      - Manufacturer generated post-market surveillance data (e.g., customer testing results);
      - Complaint handling databases; and
      - Details of clinically relevant software modifications (e.g., recalls, customer notifications, hazard alerts).
  - Conducting a scientific validity study
    - These methods of establishing an association is a planned, designed and purposefully conducted when a SaMD manufacturer is establishing an association of the intended SaMD output with a healthcare situation or condition or a physiological state. These studies commonly include prospective studies, observational studies, retrospective and longitudinal studies that establish the clinical association. (See section 6.3.1 for further considerations)

619 Note: Some low risk SaMD's are developed when the scientific validity of the output and the  
620 algorithm is still emerging. An example would be a software application that manages heart  
621 failure with medication compliance, diet and activity education, and that is subsequently shown  
622 to reduce hospitalization in those that use it fully. As the scientific and medical knowledge  
623 further develops, the initially established scientific validity might change and/or expand.

### 624 6.3.1 Considerations for Literature search to support scientific validity

625 Literature searches may be useful in circumstances in which the scientific validity of the SaMD  
626 is not initially apparent to the manufacturer.

- 627
- 628
- 629
- 630
- 631
- 632
- The data generated through literature searching should relate directly to the SaMD in question or earlier versions with justification as to why the data for the earlier versions are applicable (e.g. reports of clinical studies that have been performed by third parties).
  - When considering the relevance of data from literature searches, the SaMD manufacturer needs to consider the quality of the literature source and assess the differences between the published clinical studies and the intended SaMD use (e.g., device inputs, intended

---

<sup>5</sup> See GHTF SG5 /N7 Scientific Validity Determination and Performance Evaluation Section 7.2.3 Experience Gained by Routine Testing and GHTF SG5 /N2 Clinical Evaluation Section 6.2 Data Generated Through Clinical Experience for additional details related to these sources of data

633 user, patient population, intended use). Specifically such considerations should take into  
634 account the:

- 635 ○ Severity, disease prevalence, and natural history of the healthcare situation or
- 636 condition being diagnosed or treated;
- 637 ○ Intended target population;
- 638 ○ Intended users; and
- 639 ○ Availability of alternative diagnostic tests and current standard of care.

- 640 ● The scientific validity evidence cited in literature can provide the manufacturer in
- 641 establishing acceptable clinical performance for a SaMD.

642 See GHTF SG5 /N7 Scientific Validity Determination and Performance Evaluation Section 7.2.2  
643 Literature and GHTF SG5 /N2 Clinical Evaluation Section 6.1 Data Generated Through  
644 Literature Searching for additional details related to literature searches.

### 645 6.3.2 Considerations for Scientific Validity Studies

646 This section applies to scientific validity studies carried out by or on behalf of a manufacturer  
647 specifically for the purposes of conformity assessment in accordance with applicable regulations.  
648 Such studies are generally expected to be designed, conducted and reported in accordance and in  
649 compliance with local regulations and guidance.

650 Scientific validity studies are studies carried out by or on behalf of a manufacturer specifically  
651 for the purpose of demonstrating the safety, effectiveness and performance of the SaMD. SaMD  
652 with little or no relevant literature or clinical experience may require observational studies to  
653 validate the SaMD algorithm and demonstrate applicability to the target patient population.  
654 Observational studies are studies in which test results obtained during the study are not used for  
655 patient management and do not impact treatment decisions. The design of studies needs to be  
656 created to minimize bias and confounding and be risk-based. The design types for these studies  
657 include:

- 658 ● Cross-sectional studies where correlation of test results to the clinical condition are
- 659 established at a single point of time. In some cases, testing is performed at the initial time
- 660 point, but patients are evaluated at later time points (e.g., the SaMD is used to evaluate
- 661 the likelihood of future states, or there exists no applicable method to establish the
- 662 clinical state at the time of testing);
- 663 ● Longitudinal studies involve multiple patient measurements with the same SaMD over
- 664 time to validate the clinical performance of the SaMD;
- 665 ● Retrospective studies where the condition of the patient and the clinical association of the
- 666 output of the SaMD is known;
- 667 ● Retrospective multi-clinician multi-case studies where multiple clinicians evaluate each
- 668 case, which allows clinician variability to be taken into account;
- 669 ● Prospective studies where the SaMD is tested during the study. In the case the SaMD is
- 670 used for the determination of a patient's future state, the study will often be based on a
- 671 prospective design; and
- 672 ● Prospective-retrospective studies where the clinical status is known but the clinical
- 673 association of the output of the SaMD is established during the study. As a prospective-

674 retrospective study will use test data that was previously generated, the manufacturer  
675 should ensure that the data is segregated to ensure there is no confounding or bias by  
676 other test results.

677 See GHTF SG5 /N7 Scientific Validity Determination and Performance Evaluation Section 7.2.1  
678 Clinical Performance Studies, GHTF SG5 /N8 Clinical Performance Studies for In Vitro  
679 Diagnostic Medical Devices for additional details related to these studies.

680 NOTE: testing performed as part of the software development cycle verification and validation  
681 activities (customer feedback from focus groups, external analytical validity studies, and  
682 research studies) is not considered a clinical performance study.

#### 683 **6.4 Generating Analytical Validity Evidence for a SaMD**

684 Analytical validity evidence of a SaMD is generated during the verification and validation  
685 activities in a manufacturer's quality management system process and is always expected for a  
686 SaMD.

687 For more details refer to [SaMD N23](#).

688 Verification and validation activities to determine analytical validity for accuracy of the SaMD  
689 should consider one or more of the following:

- 690 • Algorithms described in a recognized standard (e.g., any well-known clinical assessment,  
691 method, procedure, intervention or measurement of known validity and reliability which  
692 is generally taken to be the best available, against which new tests or results and  
693 protocols are compared) that exists in literature or current standard of care (e.g., insulin  
694 dosing for a given blood glucose level);
- 695 • Comparison with a reference standard (e.g., reference standard for the detection of focal  
696 lung disease in computer aided diagnosis);
- 697 • Comparison with reference material (e.g., Coumadin<sup>6</sup> dosing for a given International  
698 Normalized Ratio (INR)); and
- 699 • Comparison to another device or SaMD that have similar association of the output to the  
700 clinical condition.

701 The use of reference databases in verification and validation activities to show analytical validity  
702 should be qualified. In addition training data sets used during the development of the SaMD  
703 algorithm should be kept separate and independent from the data set used to generate analytical  
704 validity.

705 Where the above described methods are not readily available, it may be possible to perform a  
706 comparison with an already available SaMD or a comparison to a recognized method.

707 Where there are no comparative approaches that can be used, then different approaches can be  
708 used such as comparison to a well-documented method, or comparison to a composite reference

---

<sup>6</sup> Coumadin is an anticoagulant normally used in the prevention of thrombosis and thromboembolism.

709 method. If using a composite reference, then assurances must be provided that the reference  
710 remains accurate if the parts of the composite are readjusted.

## 711 **6.5 Generating Clinical Performance Evidence for a SaMD**

712 In addition to analytical validity evidence, clinical performance evidence should be generated  
713 using process and activities that are planned, designed, conducted, analyzed and evaluated so that  
714 the best possible representation is achieved with the target population in accordance with the  
715 intended use. Optimal design, execution and analysis of such evaluation will ensure the greatest  
716 possible generalization of results (e.g., for different demographic or ethnic groups, multiple sites  
717 in different health care and geographical settings).

718 In most circumstances, clinical performance for SaMD can be generated using real or simulated  
719 data sets (e.g., automated segmentation of retinal vessels is a generally well understood problem,  
720 aided by the public availability of the annotated STARE (Structured Analysis of the Retina) and  
721 DRIVE (Digital Retinal Images for Vessel Extraction) datasets with hundreds of papers  
722 published<sup>7</sup>) that reflect real patient conditions. The SaMD manufacturer is responsible for  
723 identifying relevant data and determining the types and amount of data needed to establish  
724 clinical performance, and considering the advantages and limitations of each data type. Data  
725 relevant to the clinical performance of a SaMD may be held by the manufacturer (e.g., studies  
726 sponsored by the SaMD manufacturer) or in scientific literature (e.g., published articles of  
727 clinical performance studies related to the use of SaMD algorithms for intended clinical  
728 conditions.)

729 Before proceeding to validate the clinical performance of the SaMD in question, the  
730 manufacturer should consider:

- 731 1. Is there published clinical performance data that is not in possession of the manufacturer  
732 that may assist the manufacturer in establishing acceptable clinical performance of the  
733 SaMD?
- 734 2. Are there types of performance data available that are generated in real world use  
735 conditions that are outside the conduct of clinical performance studies?
  - 736 ○ The value of such data is that it provides real world experience obtained in larger,  
737 heterogeneous and more complex SaMD use scenarios. This type of data is also  
738 most useful for identifying less common but potentially serious device-related  
739 adverse events. It is also a particularly useful source for low risk SaMD that are  
740 based on long standing, well-characterized inputs, algorithms and outputs.
- 741 3. Are there existing SaMD or devices that have shown clinical performance for a similar  
742 association of the SaMD output to the clinical condition?
  - 743 ○ The manufacturer should determine clinical performance on both the reference  
744 device/software and the SaMD against a source of truth (i.e., gold standard) used  
745 by the original device. For example, if you were developing a software tool for

---

<sup>7</sup> Abramoff, M. D., Garvin, M. K., & Sonka, M. (2010). Retinal Imaging and Image Analysis. *IEEE Transactions on Medical Imaging*, 3, 169–208. <http://doi.org/10.1109/RBME.2010.2084567>



746 identifying a heart murmur based on an electronic stethoscope input, there may  
747 not be a way to evaluate the clinical performance of that tool only against an  
748 existing murmur detection software package. Rather, you would test both the old  
749 and new software tools against echocardiography as the reference method.

750 When selecting information for clinical performance it should be evaluated to determine its  
751 relevance and quality to address questions about the SaMD, and its contribution to demonstrate  
752 the clinical performance of the SaMD (including any specific claims about performance).

- 753 • To be relevant the information source should be specific to the SaMD in question and  
754 reflect its intended use;
- 755 • The information provided should be of sufficient quality to enable a rational and  
756 objective assessment of the clinical performance of the SaMD;
- 757 • The different data sets should be reviewed for consistency of results across multiple  
758 studies and as appropriate, the intended target populations of the SaMD;
- 759 • If the different data sets report comparable performance characteristics, certainty about  
760 the clinical performance increases. If different results are observed across the data sets, it  
761 will be helpful to determine the reason for such differences. Regardless, all data sets  
762 relevant to the SaMD should be included;
- 763 • Any risks associated with the use of the SaMD are acceptable when weighed against the  
764 benefits to the patient.

765 For novel SaMD that have no known scientific validity it may be important to generate clinical  
766 performance evidence by conducting a clinical performance study (see section **Error! Reference**  
767 **source not found.** for details.) Clinical performance studies do not necessarily imply  
768 “prospective randomized controlled trials”. Rather, depending on the risk profile of the SaMD,  
769 data (see Section 7.2) may be collected by conducting an “observational study” which is usually  
770 performed in parallel with the use of an existing SaMD, routine diagnostic testing performed for  
771 patient management care, passively collecting data while using medical devices, or in general  
772 patient care. However, for SaMD intended to diagnose or treat a healthcare situation or condition  
773 where there is a high patient risk (see Section 8.1) for inaccurate results, the study should  
774 manage risks associated and remove any bias or other confounding assumptions.

775 The following sections highlight aspects of current GHTF guidance that can be applied by taking  
776 into consideration the unique aspects of SaMD. Readers are encouraged to rely on principles and  
777 expectations in the GHTF guidance.

## 778 **6.6 Appraisal of Clinical Evaluation Evidence**

779 The SaMD manufacturer and the user(s) of SaMD should be able to reach the following  
780 conclusions through clinical evaluation: the SaMD is appropriate for its intended use; the SaMD  
781 achieves the expected performance for its intended use; the safety<sup>8</sup>, effectiveness and

---

<sup>8</sup> For more information on the concept of safety refer to [GHTF SG1 /N68 Essential Principles of Safety and Performance of Medical Devices](#) Section 5, and 6 and 7 as appropriate for software. It should be noted that the assessment of the safety of a SaMD may require more than an assessment of the clinical evaluation of the SaMD.

782 performance of the SaMD are supported by sufficient evidence; and the SaMD risks<sup>9</sup> are  
783 acceptable balanced with expected benefits<sup>10</sup>. This appraisal should consider:

- 784 • Matching SaMD intended use to the clinical evaluation evidence; and the
- 785 • Benefits and risk of the SaMD; which includes:
  - 786 • Objective consideration of patient preference in the use of the SaMD; and
  - 787 • Benefits as compared to current standard of care for the disease or condition.

788 The purpose of the appraisal of the evidence is to select information based on its merits and  
789 limitations to demonstrate that the clinical evaluation evidence matches the SaMD's intended use  
790 and related claims.

791 Each piece of information should be appraised to determine its relevance and quality. To be  
792 relevant, the information should show a clear link between the output of the SaMD to its  
793 intended use as stated in the SaMD definition statement, namely its relationship to the healthcare  
794 decision and healthcare situation or condition intended by the SaMD. The information provided  
795 should be of sufficient quality to enable a rational and objective assessment of the certainty with  
796 which the clinical evaluation evidence matches the intended use of the SaMD. It is expected that  
797 SaMD manufacturers (and third parties as appropriate) appraise the evidence generated by the  
798 clinical evaluation.

799 Specifically, appraisal of the evidence generated from clinical evaluation should address the  
800 relevance and quality of all SaMD aspects including the following:

- 801 • SaMD definition statement;
- 802 • Risk assessment and associated documentation;
- 803 • Labelling including claims, warning, limitations, contraindications, etc.;
- 804 • SaMD requirements in the QMS system; and
- 805 • Verification and validation.

#### 806 6.6.1 Matching Clinical Evaluation Evidence to SaMD's Intended Use and Related Claims

807 When the clinical evaluation evidence isn't adequate for the intended use and claims of the  
808 SaMD, it may be necessary to modify the intended use and claims to mitigate or prevent the risk  
809 of incorrect results harming patients, and to provide users with confidence in the SaMD. There  
810 should be adequate transparency to the users on the clinical validity and any limitations on the  
811 SaMD's intended use by providing appropriate contraindications, precautions, and warnings to  
812 the users in the SaMD's labeling.

---

<sup>9</sup> SaMD risks include the risk of an intervention or an unnecessary intervention or the consequences of failing to intervene as a result of inaccurate or incorrect output from a SaMD.

<sup>10</sup> For more information on the concept of safety and benefit/risk refer to [GHTF SG1 /N68 Essential Principles of Safety and Performance of Medical Devices](#) Section 6 as appropriate for software.

813 In cases where data sets used for generating clinical performance evidence are of limited  
814 availability or do not cover the desired range of the algorithm, or outputs, limitations of  
815 performance should be made transparent to the user and patients as part of the labelling.  
816 Alternatively, altering the original intended use statement and claims to match the actual  
817 performance is also considered to be adequate. For example – a novel SaMD that intends to  
818 diagnose patients with a certain condition, finds out that there is limited evidence on the  
819 acceptable analytical validity measures (accuracy, limit of detection, precision, etc). Clearly  
820 indicating in a transparent manner such actual performance is needed for the user to have  
821 confidence in the output of the SaMD, and to minimize risk to patients from inadequate results of  
822 the SaMD.

823 As stated in Section 7.2, there is flexibility regarding the type of clinical performance evidence  
824 required to establish validation of a SaMD claim(s).

#### 825 6.6.2 Benefit/Risk Determination

826 Benefit/risk determination should incorporate evidence and knowledge from the assessment of  
827 scientific validity, analytical validity, and clinical performance, but also considerations for  
828 patient preferences and alternative methods for standard of care associated with the healthcare  
829 situation or condition that the SaMD operates in. The risk tolerance varies among patients and  
830 affects the individual patients' decisions and willingness to accept such risk with the SaMD in  
831 exchange for the benefit. The assessment should focus on relevant facts, uncertainties, and key  
832 areas of judgment.

833 SaMD generally only poses risks associated with decisions made based on the output provided  
834 by the SaMD. In cases of a false positive output by the SaMD, an unnecessary test or procedure  
835 may occur, resulting in associated procedural risks, the most serious of which may include  
836 deterioration of the patient's healthcare situation or condition, the need for surgical intervention,  
837 and death. In cases of a false negative, there is risk of failure to diagnose and properly treat a  
838 significant situation or condition, which could also be associated with the same adverse events  
839 mentioned above.

840 The probable benefits of the SaMD are also based on the output provided by the SaMD. These  
841 include improved sensitivity and specificity for detecting the healthcare situation or condition  
842 compared to other available methods of care. The benefit/risk assessment should determine and  
843 evaluate the likelihood of false positives and false negatives in the intended use population and  
844 where possible compare these to known standards for sensitivity and specificity of the condition  
845 being evaluated.

846 Other methods available for accomplishing the intended use of the SaMD should be identified.  
847 The results of the evidence should indicate that the SaMD performs favorably compared to other  
848 available technologies. The benefit/risk determination should consider the impact of results that  
849 cannot be generalized to a broader population than that studied. Patients may be willing to accept  
850 the risks associated with the SaMD because of its noninvasive nature.

851 In conclusion, the available information should support the quantitative and qualitative analysis  
852 of the SaMD results for the intended use, and demonstrate that the probable benefits outweigh  
853 the probable risks for the SaMD.

854

855 **7.0 Level of Evidence According to SaMD Category**856 **7.1 Categories of SaMD**

857 *SaMD N12* describes an approach to categorize SaMD based on the factors identified in the  
 858 SaMD definition statement. The determination of the categories is the combination of the  
 859 significance of the information provided by the SaMD to the healthcare decision and the impact  
 860 of the information provided by the SaMD to the healthcare situation or condition as shown in the  
 861 table below and in Section 8.3.

State of Healthcare situation or condition	Significance of information provided by SaMD to healthcare decision		
	Treat or diagnose	Drive clinical management	Inform clinical management
Critical	IV.i	III.i	II.i
Serious	III.ii	II.ii	I.ii
Non-serious	II.iii	I.iii	I.i

862

Figure - SaMD Categories

863 The four categories (I, II, III, and IV) are based on the levels of impact on the patient or public  
 864 health where accurate information provided by the SaMD to treat or diagnose, drive clinical  
 865 management or inform clinical management is vital to avoid death, long-term disability or other  
 866 serious deterioration of health.

867 **7.2 Importance of Clinical Evidence and Expectations by SaMD Category**

868 As described previously, clinical evaluation evidence is generated to show adequate analytical  
 869 validity along with clinical validity; the level of evidence should be risk based.

870

871 The following factors are used to determine the level of clinical evaluation evidence and where  
 872 needed:

- 873 • The category of the SaMD<sup>11</sup> - Category I and Category II SaMD are considered lower  
 874 risk compared to higher risk SaMD in Categories III and IV as the latter include SaMD  
 875 that provide a diagnosis or recommendation for treatment for critical and serious  
 876 situations or conditions; and
- 877 • The intended use of the output of the SaMD - As identified in Section 6.1, SaMD can  
 878 treat a situation or condition, or can be grouped as either non-diagnostic SaMD or  
 879 diagnostic SaMD.

---

<sup>11</sup> See Appendix 8.3 – SaMD Categorization.

880 The following summarizes where clinical evaluation evidence is needed to demonstrate the  
 881 clinical evaluation of the SaMD based on the clinical evaluation that was performed using the  
 882 above factors, and based on the impact of the SaMD’s output to patients and public health:

883 **SaMD in Category I:**

- 884 • For all SaMD in this category:
  - 885 • Analytical validity evidence (generated through
  - 886 verification and validation QMS activity) based on and
  - 887 in conjunction with scientific validity information is
  - 888 sufficient to demonstrate the clinical evaluation evidence
  - 889 of the SaMD.
- 890 • For Novel SaMD in this category:
  - 891 • Manufacturers are expected to collect real world performance data to generate
  - 892 scientific validity evidence in addition to analytical validity evidence (generated
  - 893 through verification and validation QMS activity).

State of Healthcare situation or condition	Significance of information provided by SaMD to healthcare decision		
	Treat or diagnose	Drive clinical management	Inform clinical management
Critical	IV.i	III.i	II.i
Serious	III.ii	II.ii	I.ii
Non-serious	II.iii	I.iii	I.i

894 **SaMD in Category II**

- 895 • For all SaMD except for category II.ii.,
  - 896 • Analytical validity evidence (generated through
  - 897 verification and validation QMS activity) based on and
  - 898 in conjunction with scientific validity information is
  - 899 sufficient to demonstrate the clinical evaluation
  - 900 evidence of the SaMD.
- 901 • For Diagnostic SaMD in II.iii:
  - 902 • Clinical performance evidence is expected in addition to analytical validity and
  - 903 scientific validity evidence.
- 904 • For Novel SaMD in this category:
  - 905 • Manufacturers are expected to collect real world performance data to generate
  - 906 scientific validity evidence in addition to analytical validity evidence (generated
  - 907 through verification and validation QMS activity).

State of Healthcare situation or condition	Significance of information provided by SaMD to healthcare decision		
	Treat or diagnose	Drive clinical management	Inform clinical management
Critical	IV.i	III.i	II.i
Serious	III.ii	II.ii	I.ii
Non-serious	II.iii	I.iii	I.i

908 **SaMD in Categories II.ii, III and IV:**

- 909 • For all SaMD in these categories (well-known or novel):
  - 910 • Analytical validity evidence (generated through
  - 911 verification and validation QMS activity) based on and
  - 912 in conjunction with scientific validity information is
  - 913 sufficient to demonstrate the clinical evaluation evidence
  - 914 of the SaMD.
  - 915 • In circumstances where the scientific validity is novel, manufacturers should generate
  - 916 appropriate association of the SaMD output to the clinical condition/physiological
  - 917 state using approaches described in scientific validity as described in Section 6.3.
- 918 • For Diagnostic SaMD in these categories:
  - 919 • Clinical performance evidence is expected in addition to analytical validity and
  - 920 scientific validity evidence.

State of Healthcare situation or condition	Significance of information provided by SaMD to healthcare decision		
	Treat or diagnose	Drive clinical management	Inform clinical management
Critical	IV.i	III.i	II.i
Serious	III.ii	II.ii	I.ii
Non-serious	II.iii	I.iii	I.i

Legend:					
<p><b>Non-D<sub>x</sub>-SaMD</b> = Treat / Non-Diagnostic SaMD</p> <p><b>D<sub>x</sub>-SaMD</b> = Diagnostic SaMD</p> <p><b>AV + SV</b> = Analytical validity + Scientific Validity</p> <p><b>AV + SV + CP</b> = Analytical validity + Scientific Validity + Clinical Performance</p>			<ul style="list-style-type: none"> <li>Treat;</li> <li>Provide therapy to a human body using other means;</li> <li>Diagnose;</li> <li>Detect;</li> <li>Screen;</li> <li>Prevent;</li> <li>Mitigate;</li> <li>Lead to an immediate or near term action.</li> </ul>	<ul style="list-style-type: none"> <li>Aid in treatment;</li> <li>Provide enhanced support to safe and effective use of medicinal products;</li> <li>Aid in diagnosis;</li> <li>Help predict risk of a disease or condition;</li> <li>Aid to making a definitive diagnosis;</li> <li>Triage early signs of a disease or condition;</li> <li>Identify early signs of a disease or condition.</li> </ul>	<ul style="list-style-type: none"> <li>Inform of options for treatment;</li> <li>Inform of options for diagnosis;</li> <li>Inform of options for prevention;</li> <li>Aggregate relevant clinical information;</li> <li>Will not trigger an immediate or near term action.</li> </ul>
			<b>Treat or Diagnose</b>	<b>Drive Clinical Management</b>	<b>Inform Clinical Management</b>
<ul style="list-style-type: none"> <li>Life-threatening;</li> <li>Fragile</li> </ul>	<ul style="list-style-type: none"> <li>Requires major therapeutic interventions;</li> <li>Sometimes time critical</li> <li>Vital to: avoiding death; serious deterioration of health; mitigating public health situations or conditions</li> </ul>	<ul style="list-style-type: none"> <li>Specialized trained users</li> </ul>	<b>Critical</b>	<p><b>TYPE IV.i</b></p> <p><b>TYPE III.i</b></p> <p><b>Independent Review is important</b></p> <p>Non-D<sub>x</sub>-SaMD → AV + SV</p> <p>D<sub>x</sub>-SaMD → AV + SV + CP</p>	<b>TYPE II.i</b>
<ul style="list-style-type: none"> <li>Moderate in progression</li> <li>Often curable;</li> <li>Not fragile;</li> </ul>	<ul style="list-style-type: none"> <li>Does not require major therapeutic interventions</li> <li>Not expected to be time critical</li> <li>Vital to avoiding unnecessary interventions</li> </ul>	<ul style="list-style-type: none"> <li>Either specialized trained users or lay users.</li> </ul>	<b>Serious</b>	<p><b>TYPE III.ii</b></p> <p><b>TYPE II.ii</b></p>	<b>TYPE I.ii</b>
<ul style="list-style-type: none"> <li>Slow with predictable progression of disease state</li> <li>Minor chronic illnesses or states</li> <li>May not be curable;</li> <li>Individuals who may not always be patients</li> <li>Can be managed effectively</li> </ul>		<ul style="list-style-type: none"> <li>Either specialized trained users or lay users</li> </ul>	<b>Non-Serious</b>	<p><b>TYPE II.iii</b></p> <p><b>TYPE I.iii</b></p> <p><b>Document AV, SV and CP -- Independent Review not important</b> (For Novel SaMD – Build SV and CP evidence using “Real World” experience)</p> <p>Non-D<sub>x</sub>-SaMD → AV + SV</p> <p>D<sub>x</sub>-SaMD → AV + SV + CP</p>	<b>TYPE I.i</b>
<b>Disease Type /Patient Condition</b>	<b>Intervention Type</b>	<b>User Type</b>			

921  
922 **Figure 8 - Summary of Clinical Evidence and Expectations by SaMD Category (See appendix 8.5 for full page image)**

923 **7.3 Importance of Independent Review of Evidence by SaMD Category**

924 Similar to the importance of evidence, certain SaMD categories may require independent review  
925 of the evidence to provide users the confidence in the SaMD’s clinical validity. The concept of  
926 independent review is analogous to having peer review of journal articles or the concept of  
927 design review performed in the QMS system.

928 The recommendation for independent review for certain categories of SaMD does not imply the  
929 need for premarket review (authorization) by a regulatory authority which is outside the scope of  
930 this document. Regardless of the category of SaMD, the level of regulatory oversight (premarket  
931 review/market authorization) may depend on an individual jurisdiction’s regulatory laws where  
932 the SaMD will be made available.

933 The recommendation for independent review highlights where the evidence generated from the  
934 clinical evaluation of the SaMD should be reviewed by someone other than the SaMD  
935 manufacturer to objectively appraise the SaMD’s intended purpose and the conformity with the  
936 overall clinical evaluation evidence.



937 The following is a possible recommendation where independent review of clinical evaluation  
 938 evidence is of importance.

939 **SaMD in Category I:**

- 940 • Independent review of evidence not important
- 941 • Manufacturers should document their appraisal of the
- 942 clinical evaluation evidence with the SaMD definition
- 943 statement and associated claims.

State of Healthcare situation or condition	Significance of information provided by SaMD to healthcare decision		
	Treat or diagnose	Drive clinical management	Inform clinical management
Critical	IV.i	III.i	II.i
Serious	III.ii	II.ii	L.ii
Non-serious	II.iii	L.iii	L.i

944 **SaMD in Category II (except for category II.ii):**

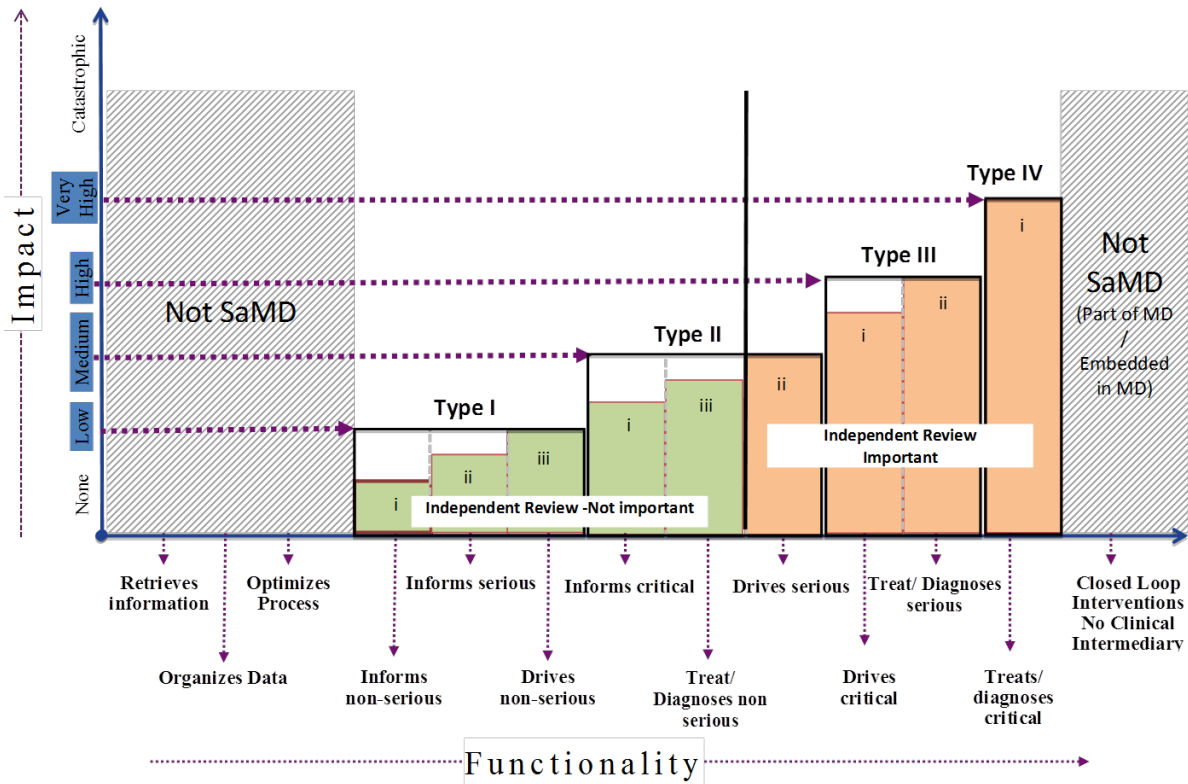
- 945 • Independent review of evidence not important
- 946 • Manufacturers should document their appraisal of the
- 947 clinical evaluation evidence with the SaMD definition
- 948 statement and associated claims.

State of Healthcare situation or condition	Significance of information provided by SaMD to healthcare decision		
	Treat or diagnose	Drive clinical management	Inform clinical management
Critical	IV.i	III.i	II.i
Serious	III.ii	II.ii	L.ii
Non-serious	II.iii	L.iii	L.i

949 **SaMD in Categories II.ii, III and IV:**

- 950 • Manufacturers should document their appraisal of the
- 951 clinical evaluation evidence with the SaMD definition
- 952 statement and associated claims.
- 953 • Independent review of evidence is important
- 954

State of Healthcare situation or condition	Significance of information provided by SaMD to healthcare decision		
	Treat or diagnose	Drive clinical management	Inform clinical management
Critical	IV.i	III.i	II.i
Serious	III.ii	II.ii	L.ii
Non-serious	II.iii	L.iii	L.i



955  
 956  
 957

Figure 9: Importance independent review

---

## 958 7.4 Pathway for Continuous Learning Leveraging Real World Clinical Evidence

959 It is anticipated that one of the unique aspects that differentiate SaMD from other medical  
960 devices is the way SaMD may leverage technology and connectivity i.e., the seamless  
961 communication between devices, technology and people to continuously monitor the safety,  
962 effectiveness and performance of the SaMD. Unlike many other medical devices where real  
963 world experience is often difficult to gather as it comes in many forms (e.g., longitudinal follow  
964 up data that may be in a registry or insurance claims) and quality (e.g., missing data, variable  
965 definitions, etc.), with the connectivity of a SaMD this is easier.

966 Ideally the SaMD manufacturer has an idea early on regarding the longer term possibilities for  
967 the functionality and claims that may be supported by learning about the SaMD over time. As  
968 additional clinical data to support the new claims is gathered, the SaMD manufacturer will  
969 update the clinical evaluation. In practice, the clinical evaluation is a dynamic summary that  
970 changes as knowledge of the SaMD increases.

971 The “continuous learning” referred to here is not ‘machine learning software’, i.e., where  
972 software device keeps learning automatically after it has been released into the market; rather it  
973 refers to collecting post-market information.

974 Continuously collecting and analyzing post-market information (e.g., safety reports, including  
975 adverse event reports, results from performance studies, published literature) can help the SaMD  
976 manufacturer understand the real world performance of the SaMD. Manufacturers should  
977 appropriately review this information to determine if there are any changes to the safety,  
978 effectiveness or performance, or possible impact on benefits and risks of the SaMD that would  
979 indicate a need for a design change or a labeling change regarding contraindications, warnings,  
980 precautions or instructions for use.

981 It is also anticipated that if planned correctly, as a SaMD manufacturer learns by monitoring real  
982 world experience it can help the SaMD evolve after introduction into the market. This may  
983 potentially lead to a substantial change to the SaMD intended use and claims supported by the  
984 clinical data gathered, analyzed and appraised from the continuous monitoring.

985 Learning may impact the original category of a SaMD in the following ways:

- 986 • Real world performance provides evidence that analytical or clinical performance is  
987 superior than the performance initially evaluated by SaMD manufacturer, or
- 988 • Real world evidence indicates that analytical or clinical performance is lower than the  
989 performance initially evaluated by SaMD manufacturer.

990 An example is shown in scenario 1 in Figure 10 below. In this scenario, a SaMD manufacturer  
991 can conduct a retrospective clinical evaluation based on real world data and incorporate new  
992 information into the SaMD claims to enhance its clinical validity by further clarifying the  
993 SaMD’s performance.

994 In the example shown as scenario 2 in Figure 10 below, a SaMD manufacturer can conduct a  
995 clinical evaluation based on gathering prospective real world data and incorporating the new  
996 information into the SaMD’s intended use and definition statement, modifying design features to  
997 minimize risk, provide transparency by further clarifying the SaMD’s performance and validity,



998 and minimize risk of incorrect results resulting in patient harm. Such data can potentially result  
 999 in modification of the impact (risk) category of a SaMD from high to medium.

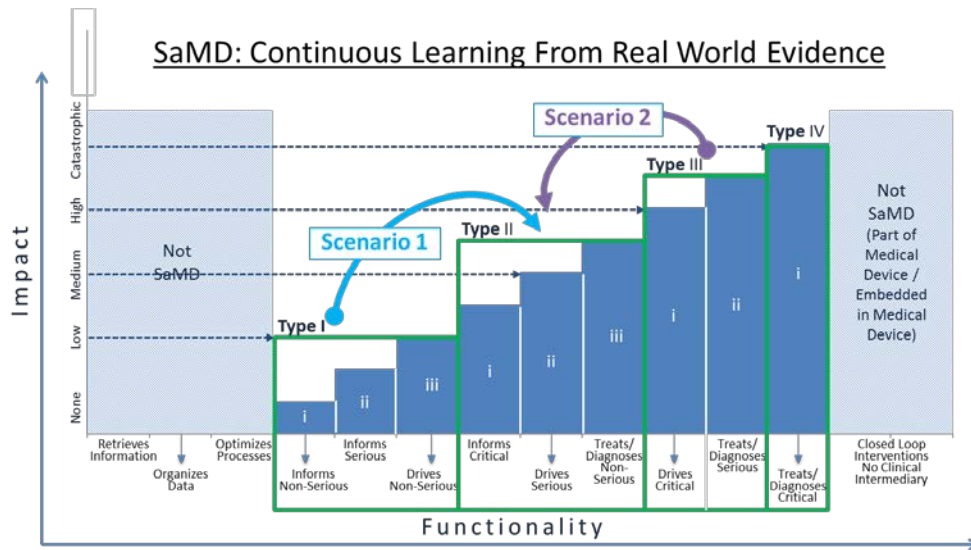


Figure 10: Continuous learning from real world evidence

1000  
1001

1002 This document encourages SaMD manufacturers to leverage SaMD’s unique capability to  
 1003 capture user’s interactions with the SaMD to conduct well planned clinical performance  
 1004 observational studies in addition to ongoing monitoring of technical and clinical performance.

1005 A SaMD manufacturer can conduct an observational study that takes into consideration the  
 1006 healthcare situation or condition, and support a higher level significance of the information. For  
 1007 example, the output of a SaMD that is initially in the market to “inform” a serious healthcare  
 1008 situation or condition can collect evidence and provide the input data set to support claims for the  
 1009 output of the SaMD to either “drive” or “diagnose” a serious healthcare situation or condition. It  
 1010 would be expected that when moving up in significance from “inform” to either “drive” or  
 1011 “diagnose”, that the same rigor be applied in evaluating scientific validity, analytical validity and  
 1012 clinical performance where appropriate as recommended in Section 7.3 . The advantage for the  
 1013 SaMD manufacturer is that they would access the data set that can support the evaluation with  
 1014 real world observational data and a retrospective analysis.

1015 To summarize, one can envision a “building block” approach or an agile clinical evidence  
 1016 gathering approach to assimilating clinical evidence for a SaMD based on its risk categorization.  
 1017 Risk categorization of the SaMD is an evolving phenomenon through the lifecycle of the SaMD  
 1018 based on the on-going clinical evaluation process for the SaMD. All modifications that result  
 1019 from real world experience should also follow the framework for evidence requirements as  
 1020 outlined in Section 7.2 and level of independent review as outlined in Section 7.3.

1021

---

## 1022 8.0 Appendices

### 1023 8.1 SaMD Definition Statement

1024 All manufacturers should, as highlighted below and in ( Section 6.0, start with a SaMD definition  
1025 statement that is clear and strong about the intended use of the SaMD. Generally these aspects  
1026 can be grouped into the following two major factors that provide adequate description of the  
1027 intended use of SaMD:

- 1028 **A. The “significance of the information provided by the SaMD to the healthcare**  
1029 **decision”** which identifies the intended medical purpose of the SaMD. The statement  
1030 should explain how the SaMD meets one or more of the purposes described in the  
1031 definition of a medical device, e.g., supplying information for diagnosis, prevention,  
1032 monitoring, treatment etc. structured in following sub categories:
- 1033 a. **To treat or to diagnose** – the information provided by the SaMD will be used to take  
1034 an immediate or near term action:
- 1035 i. To treat/prevent or mitigate by connecting to other medical devices, medicinal  
1036 products, general purpose actuators or other means of providing therapy to a  
1037 human body
- 1038 ii. To diagnose/screen/detect a disease or condition (i.e., using sensors, data, or  
1039 other information from other hardware or software devices, pertaining to a  
1040 disease or condition)
- 1041 b. **To drive clinical management** - the information provided by the SaMD will be used  
1042 to aid in treatment, aid in diagnoses, to triage or identify early signs of a disease or  
1043 condition will be used to guide next diagnostics or next treatment interventions:
- 1044 i. To aid in treatment by providing enhanced support to safe and effective use of  
1045 medicinal products or a medical device.
- 1046 ii. To aid in diagnosis by analyzing relevant information to help predict risk of a  
1047 disease or condition or as an aid to making a definitive diagnosis.
- 1048 iii. To triage or identify early signs of a disease or conditions.
- 1049 c. **To Inform clinical management** – the information provided by the SaMD will not  
1050 trigger an immediate or near term action:
- 1051 i. To inform of options for treating, diagnosing, preventing, or mitigating a  
1052 disease or condition.
- 1053 ii. To provide clinical information by aggregating relevant information (e.g.,  
1054 disease, condition, drugs, medical devices, population, etc.)
- 1055 **B. The intended “state of the healthcare situation or condition”** that identifies the  
1056 intended use for a disease or condition taking into account the patient’s state of health,  
1057 progression of the disease and associated type and immediacy of interventions, target  
1058 population and type of users (trained or lay users). This portion of the statement should be  
1059 expressed in the following structured sub categories:

- 1060 a. **Critical situation or condition** - Situations or conditions where accurate and/or  
1061 timely diagnosis or treatment action is vital to avoid death, long-term disability or  
1062 other serious deterioration of health of an individual patient or to mitigating impact to  
1063 public health. SaMD is considered to be used in a critical situation or condition  
1064 where:
- 1065 i. The type of disease or condition is:
- 1066 1. Life-threatening state of health, including incurable states,  
1067 2. Requires major therapeutic interventions,  
1068 3. Sometimes time critical, depending on the progression of the disease  
1069 or condition that could affect the user's ability to reflect on the output  
1070 information.
- 1071 ii. Intended target population is fragile with respect to the disease or condition  
1072 (e.g., pediatrics, high risk population, etc.)
- 1073 iii. Intended for specialized trained users.
- 1074 b. **Serious situation or condition** - Situations or conditions where accurate diagnosis or  
1075 treatment is of vital importance to avoid unnecessary interventions (e.g., biopsy) or  
1076 timely interventions are important to mitigate long term irreversible consequences on  
1077 an individual patient's health condition or public health. SaMD is considered to be  
1078 used in a serious situation or condition when:
- 1079 i. The type of disease or condition is:
- 1080 1. Moderate in progression, often curable,  
1081 2. Does not require major therapeutic interventions,  
1082 3. Intervention is normally not expected to be time critical in order to  
1083 avoid death, long-term disability or other serious deterioration of  
1084 health, whereby providing the user an ability to detect erroneous  
1085 recommendations.
- 1086 ii. Intended target population is NOT fragile with respect to the disease or  
1087 condition.
- 1088 iii. Intended for either specialized trained users or lay users.
- 1089 Note: SaMD intended to be used by lay users in a "serious situation or condition" as  
1090 described here, without the support from specialized professionals, should be  
1091 considered as SaMD used in a "critical situation or condition".
- 1092 c. **Non-Serious situation or condition** - Situations or conditions where an accurate  
1093 diagnosis and treatment is important but not critical for interventions to mitigate long  
1094 term irreversible consequences on an individual patient's health condition or public  
1095 health. SaMD is considered to be used in a non-serious situation or condition when:
- 1096 i. The type of disease or condition is:
- 1097 1. Slow with predictable progression of disease state (may include minor  
1098 chronic illnesses or states),  
1099 2. May not be curable; can be managed effectively,  
1100 3. Requires only minor therapeutic interventions, and

- 1101 4. Interventions are normally noninvasive in nature, providing the user  
 1102 the ability to detect erroneous recommendations.  
 1103 ii. Intended target population is individuals who may not always be patients.  
 1104 iii. Intended for use by either specialized trained users or lay users.

1105 C. **Description of the SaMD’s core functionality**<sup>12</sup> which identifies the critical  
 1106 features/functions of the SaMD that are essential to the intended significance of the  
 1107 information provided by the SaMD to the healthcare decision in the intended healthcare  
 1108 situation or condition. This description should include only the critical features. (See  
 1109 applicability of this in ( Section 6.0).

1110 For more details and information related to the two major factors and formulating the SaMD  
 1111 Definition Statement refer to ( Sections 5.0 and 6.0.

## 1112 8.2 Clarifying SaMD Definition

1113 This Appendix provides a representative list of features and functionalities that either meet or  
 1114 don’t meet the definition of SaMD. This list is not exhaustive; it is only intended to provide  
 1115 clarity and assistance in identifying when a feature or functionality is considered to be SaMD.

### 1116 *Examples of software that are SaMD:*

- 1117 • Software with a medical purpose that operates on a general purpose computing platform,  
 1118 i.e., a computing platform that does not have a medical purpose, is considered SaMD. For  
 1119 example, software that is intended for diagnosis of a condition using the tri-axial  
 1120 accelerometer that operates on the embedded processor on a consumer digital camera is  
 1121 considered a SaMD.
- 1122 • Software that is connected to a hardware medical device but is not needed by that  
 1123 hardware medical device to achieve its intended medical purpose is SaMD and not an  
 1124 accessory to the hardware medical device. For example, software that allows a  
 1125 commercially available smartphone to view images for diagnostic purposes obtained  
 1126 from a magnetic resonance imaging (MRI) medical device is SaMD and not an accessory  
 1127 to MRI medical device.
- 1128 • The SaMD definition notes states that “SaMD is capable of running on general purpose (non-  
 1129 medical purpose) computing platforms.” SaMD running on these general purpose computing  
 1130 platform could be located in a hardware medical device, For example, software that  
 1131 performs image post-processing for the purpose of aiding in the detection of breast cancer  
 1132 (CAD - computer-aided detection software) running on a general purpose computing  
 1133 platform located in the image-acquisition hardware medical device is SaMD.
- 1134 • The SaMD definition notes states that “SaMD may be interfaced with other medical devices,  
 1135 including hardware medical devices and other SaMD software, as well as general purpose

---

<sup>12</sup> These could include specific functionality that is critical to maintain safety, effectiveness and performance profile attributes identified by risk management process undertaken by the manufacturer of SaMD.

1136 software.” Software that provides parameters that become the input for a different hardware  
 1137 medical device or other SaMD is SaMD. For example, treatment planning software that  
 1138 supplies information used in a linear accelerator is SaMD.

1139 **Examples of software that are not SaMD:**

- 1140 • The SaMD definition states “SaMD is defined as software intended to be used for one or  
 1141 more medical purposes that perform these purposes without being part of a hardware medical  
 1142 device”. Examples of software that are considered “part of” include software used to “drive  
 1143 or control” the motors and the pumping of medication in an infusion pump; or software used  
 1144 in closed loop control in an implantable pacemaker or other types of hardware medical  
 1145 devices. These types of software, sometimes referred to as “embedded software”,  
 1146 “firmware”, or “micro-code” are, not SaMD”.
- 1147 • Software required by a hardware medical device to perform the hardware’s medical  
 1148 device intended use is not SaMD even if/when sold separately from the hardware medical  
 1149 device.
- 1150 • Software that relies on data from a medical device, but does not have a medical purpose,  
 1151 e.g., software that encrypts data for transmission from a medical device is not SaMD.
- 1152 • Software that enables clinical communication and workflow including patient  
 1153 registration, scheduling visits, voice calling, and video calling is not SaMD.
- 1154 • Software that monitors performance or proper functioning of a device for the purpose of  
 1155 servicing the device, e.g., software that monitors X-Ray tube performance to anticipate  
 1156 the need for replacement; or software that integrates and analyzes laboratory quality  
 1157 control data to identify increased random errors or trends in calibration on IVDs is not  
 1158 SaMD.
- 1159 • Software that provides parameters that become the input for SaMD is not SaMD if it does  
 1160 not have a medical purpose. For example, a database including search and query  
 1161 functions by itself or when used by SaMD is not SaMD.

1162 **8.3 SaMD Categorization**

1163 ( describes a method for categorizing SaMD based on two major factors representing aspects that  
 1164 can raise or lower a SaMD's potential to create hazardous situations to patients:

- 1165 • State of the healthcare situation or condition; and
- 1166 • Significance of the information provided by the SaMD to the healthcare decision.

1167 With consideration of these two parameters, the table below displays SaMD categories:

State of Healthcare situation or condition	Significance of information provided by SaMD to healthcare decision		
	Treat or diagnose	Drive clinical management	Inform clinical management
Critical	IV.i	III.i	II.i

---

Serious	III.ii	II.ii	I.ii
Non-serious	II.iii	I.iii	I.i

1168

1169 **Criteria for Category IV –**

- 1170 i. SaMD that provides information to treat or diagnose a disease or conditions in a critical  
1171 situation or condition is a Category IV and is considered to be of very high impact.

1172 **Criteria for Category III –**

- 1173 i. SaMD that provides information to treat or diagnose a disease or conditions in a serious  
1174 situation or condition is a Category III and is considered to be of high impact.  
1175 ii. SaMD that provides information to drive clinical management of a disease or conditions  
1176 in a critical situation or condition is a Category III and is considered to be of high impact.

1177 **Criteria for Category II –**

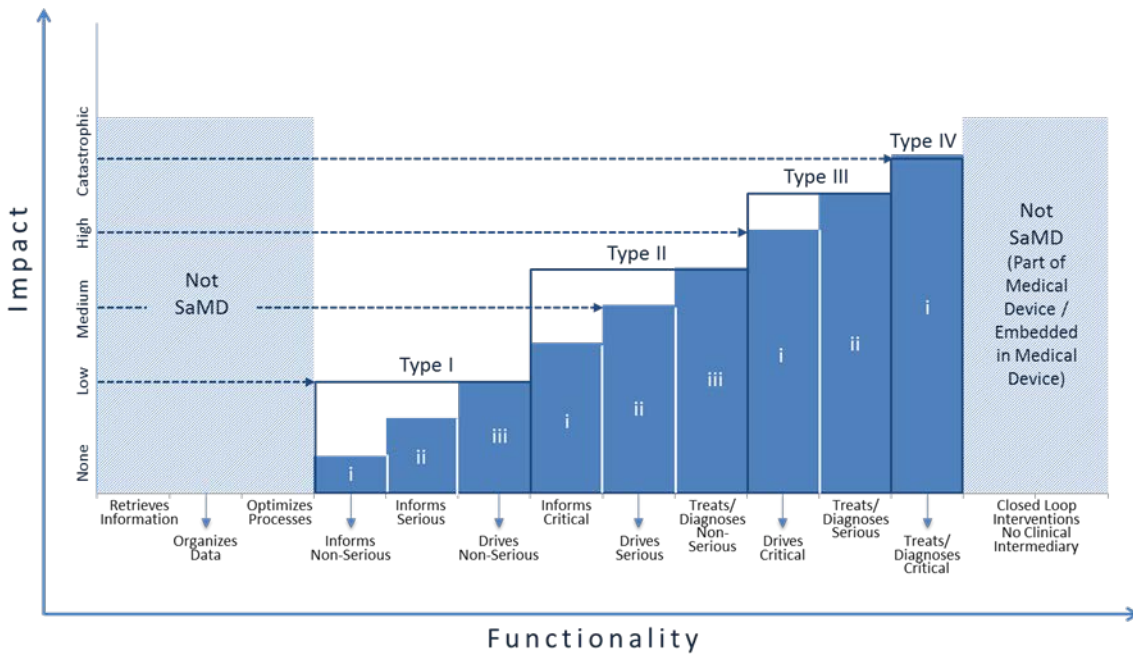
- 1178 i. SaMD that provides information to treat or diagnose a disease or conditions in a non-  
1179 serious situation or condition is a Category II and is considered to be of medium impact.  
1180 ii. SaMD that provides information to drive clinical management of a disease or conditions  
1181 in a serious situation or condition is a Category II and is considered to be of medium  
1182 impact.  
1183 iii. SaMD that provides information to inform clinical management for a disease or  
1184 conditions in a critical situation or condition is a Category II and is considered to be of  
1185 medium impact.

1186 **Criteria for Category I –**

- 1187 i. SaMD that provides information to drive clinical management of a disease or conditions  
1188 in a non-serious situation or condition is a Category I and is considered to be of low  
1189 impact.  
1190 ii. SaMD that provides information to inform clinical management for a disease or  
1191 conditions in a serious situation or condition is a Category I and is considered to be of  
1192 low impact.  
1193 iii. SaMD that provides information to inform clinical management for a disease or  
1194 conditions in a non-serious situation or condition is a Category I and is considered to be  
1195 of low impact.

1196 The figure below depicts the categories of SaMD based on the impact and functionality. As  
1197 displayed in the table above, the impact of the SaMD on patient or the public health is divided  
1198 into four categories (Categories I, II, III, IV) while functionality (to inform or drive clinical  
1199 management, to treat or diagnose) includes three categories. This categorization framework  
1200 builds on the principles underlying the classification rules established in the GHTF classification  
1201 principles documents, covering individual risks, public health risks, user skills, and importance  
1202 of the information provided. While the categorization framework itself is not a regulatory  
1203 classification, it sets a path towards a common vocabulary and approach to such classification  
1204 aimed at determining appropriate levels of regulatory oversight.

# SaMD Types Landscape / Scope



1205

1206

---

## 1207 8.4 Illustrative Examples of Clinical Evaluation Concepts for SaMD

1208 The following illustrates a series of questions for different examples that may help to determine  
1209 the required level of clinical evaluation.

### 1210 **Example: Algorithm to Detect Atrial Fibrillation**

1211 *The SaMD demonstrates with certainty (success criteria) that the algorithm is able to detect*  
1212 *atrial fibrillation with PPV of 65%.*

- 1213 • Is this a clinically adequate criterion for the intended use?
- 1214 • What are the other clinical performance specifications that are necessary in order to fully  
1215 assess this criteria (NPV, sample size, etc.)?
- 1216 • What is the population for which this detection is intended and does this have an impact  
1217 on the success criteria?
- 1218 • Does this provide a clinically meaningful outcome/result in the current standard of care?

### 1219 **Example: Algorithm interprets Myocardial Infarction**

1220 *The SaMD demonstrates with certainty (success criteria) that the algorithm can interpret*  
1221 *Myocardial Infarction with 90% accuracy.*

- 1222 • What is the sensitivity and specificity of the result?
- 1223 • How does this impact clinical workflow?
- 1224 • How does 90% accuracy fit into current standard of care or when compared to the  
1225 existing interpretation devices/SaMD?
- 1226 • What is the comparator/gold standard?
- 1227 • What is the health care situation (environment) of use and the importance of the SaMD to  
1228 clinical management?
- 1229 • What is the severity of the condition and what are the risks associated with an inaccurate  
1230 result?

### 1231 **Example: EEG Analysis**

1232 *The SaMD demonstrates with certainty that the SaMD can determine the location of a seizure*  
1233 *based on EEG?*

- 1234 • What is the scientific validity for the association of EEG signals to the location of the  
1235 seizure?
- 1236 • If no existing gold standard, what is/are the criteria for diagnosis or management and is  
1237 this clinically meaningful in the context of use for the device?
- 1238 • Did the testing results demonstrate adequate clinical performance (specificity, selectivity,  
1239 PPV, NPV, etc.)?
- 1240 • How does the availability of such SaMD output show benefits compared to current  
1241 standard of care?



---

## 1242 8.4.1 Illustrative Example of Clinical Evaluation Concepts – Skin Disorders

### 1243 **Example – Skin Disorder 1**

#### 1244 Definition Statement

1245 The SaMD provides generic information on moles, benign and atypical nevus, and malignant  
1246 skin lesions. The SaMD uses photos with rulers next to them. The user manually identifies  
1247 the location of the suspect skin lesion on a human body map, and tracks the changes over  
1248 time in terms of size and appearance. The user is prompted to seek a medical professional’s  
1249 opinion. The SaMD allows the user to send the photos to their family doctor.

1250 Based on the above definition statement the SaMD *informs clinical management*. Because  
1251 the spectrum of the skin conditions includes information related to malignant skin lesions, the  
1252 SaMD is used in a *critical healthcare situation or condition*.

1253 This is an example of a Category II.i SaMD used for non-diagnostic purposes.

#### 1254 Clinical Evaluation

1255 As a Category II.i non-diagnostic SaMD it is recommended that the manufacturer perform a  
1256 clinical evaluation providing evidence for the scientific validity and analytical validity of the  
1257 SaMD.

- 1258 • Evidence of the scientific validity may be found in literature searches and clinical  
1259 research and may include for example the use of well-known diagnostic rules in  
1260 dermatology such as the ADCDE (may also be referred to as ABCD) Rule for  
1261 mapping the mole.
- 1262 • Evidence of the analytical validity may include thoroughly checking that the results  
1263 from multiple executions of the SaMD processing the input and output satisfy the  
1264 expected or desirable properties derived from the software specification or user  
1265 expectations.

### 1266 **Example – Skin Disorder 2**

#### 1267 Definition Statement

1268 The SaMD provides lesion-specific information and flags suspect lesions that have a higher  
1269 likelihood to progress to an atypical nevus state or are clearly abnormal. The SaMD tracks  
1270 lesions with the use of color-calibrated photos of a tested minimal image quality and  
1271 promptly detects any changes to margins, size, color, reflectivity, texture, and numbers. The  
1272 SaMD automatically maps the skin lesions, highlights new lesions, counts them, and sends  
1273 photos to a dermatologist or dermatopathologist without user intervention. The SaMD drives  
1274 the next diagnostic action of a dermatologist, who’s primary goal is to decide what lesions  
1275 need interventions (excision and biopsy), and which lesions are OK to observe and monitor.

1276 Based on the above definition statement the SaMD *drives clinical management*. Because the  
1277 spectrum of the skin conditions includes information related to malignant skin lesions, the  
1278 SaMD is used in a *critical condition*.

1279 This is an example of a Category III.i SaMD used for diagnostic purposes.

---

1280 Clinical Evaluation

1281 As a Category III.i diagnostic SaMD it is recommended that the manufacturer perform a  
1282 clinical evaluation to provide evidence of clinical performance in addition to evidence for the  
1283 scientific validity and analytical validity of the SaMD.

- 1284 • Evidence of the scientific validity may be found in literature searches and clinical  
1285 research and may include for example the use of well-known diagnostic rules in  
1286 dermatology such as the ADCDE (may also be referred to as ABCD) Rule for  
1287 mapping the mole.
- 1288 • Evidence of clinical performance demonstrating that the SaMD can stratify lesions  
1289 into high and low-risk category as efficiently as a dermatologist is necessary to  
1290 demonstrate the clinical performance. This could be prospective trial or retrospective  
1291 clinical evaluation of a validated database of skin lesions (assuming the input to the  
1292 SaMD will be of the same high quality photos as found in the validated database).
- 1293 • Evidence of the analytical validity may include thoroughly checking that the results  
1294 from multiple executions of the SaMD processing the input and output satisfy the  
1295 expected or desirable properties derived from the software specification or user  
1296 expectations.

1297 For this kind of diagnostic SaMD, the clinical validity evidence that includes scientific  
1298 validation and clinical performance should be independently reviewed along with the  
1299 analytical validity evidence that will provide input to assurance of safety, effectiveness and  
1300 performance of the SaMD.

1301 **Example – Skin Disorder 3**

1302 Definition Statement

1303 The SaMD replaces the histo-pathology microscopic evaluation of a biopsy/excised sample  
1304 through the use of a high magnification lens and an external UV light source that detects  
1305 cytologic atypia (very large cells, poor maturation of cells, growth patterns) or cells typical of  
1306 malignant melanoma.

1307 Based on the above definition statement the SaMD provides a *diagnosis*. Because the  
1308 spectrum of the skin conditions includes information related to malignant skin lesions, the  
1309 SaMD is used in a *critical* condition.

1310 This is an example of a Category IV.i SaMD used for diagnostic purposes.

1311 Clinical Evaluation

1312 As a novel Category IV.i diagnostic SaMD it is recommended that the manufacturer perform  
1313 a clinical evaluation providing evidence for the scientific validity along with clinical  
1314 performance evidence to show clinical validity in addition to analytical validity of the SaMD.  
1315 such evaluation should include:

- 1316 • Evidence of the scientific validity may be found in literature searches and clinical  
1317 research that shows evidence that include using high magnification of images taken  
1318 under UV light combined with image recognition to detect malignant skin lesions

- 1319
- 1320
- 1321
- 1322
- 1323
- 1324
- 1325
- 1326
- 1327
- 1328
- Evidence of clinical performance that is generated through a study ( e.g. prospective study) comparing specificity and sensitivity of the SaMD based on histo-pathology microscopic or some genetic testing of excised lesions to confirm the diagnosis. Such study should include considerations for removing skin color, ambient light, contrast and other biases that show definitively the detection of malignant lesions. This may also require an adequate follow-up of lesions not excised/biopsied to confirm patient outcomes. There may be a need to consider that some cases may not present with skin lesions, but metastatic disease.
  - Further real world experience from user feedback should be gathered post-market on an ongoing basis to continue to evaluate the SaMD’s clinical performance.

1329 Alternative claims and additional considerations

1330 The above examples either specifically address melanoma or melanoma is within the  
1331 spectrum of the claims.

- 1332
- 1333
- 1334
- 1335
- 1336
- 1337
- 1338
- If the SaMD claims that it intends to detect furuncles, burns, frostbite, psoriasis, neurofibromatosis, chickenpox skin lesions, etc. the SaMD would be intended to be used in a serious situation or condition rather than intended to be used for a critical situation or condition thus lowering the risk profile of the SaMD.
  - If the SaMD claims to detect benign skin lesions, such as eczema, acne, cellulitis, keloids, warts, etc. – the SaMD would be used in a non-serious situation or condition lowering the risk profile of the SaMD even further.

1339 An example of scientific validity and acceptable “reference standard” for clinical  
1340 performance includes an agreement between dermatopathologists reading histology slides  
1341 under microscope. According to identified studies, there is only 35-58% concordance for  
1342 grading of dysplasia (Duncan 1993), and dermatopathologists often did not agree with  
1343 their own assessment of the same slide 6 months later (Piepkorn 1994); there is only 33%  
1344 agreement on all benign versus all malignant in a sample of 37 "clear-cut" cases (Farmer,  
1345 1996).

1346 **Example – Coronary Physiological Simulation Software**

1347 Definition Statement

1348 The software provides simulated functional assessment of blood flow in the coronary  
1349 vascular system using data extracted from medical device imaging to solve algorithms and  
1350 yield simulated metrics of physiological information (e.g., blood flow, coronary flow reserve,  
1351 fractional flow reserve, myocardial perfusion). The SaMD is intended to generate results for  
1352 use and review by a qualified clinician. This is a post-processing software for the clinical  
1353 quantitative and qualitative analysis of previously acquired Computed Tomography (CT)  
1354 DICOM<sup>13</sup> data for clinically stable symptomatic patients with coronary artery disease. The  
1355 software displays the coronary anatomy with functional information using graphics and text,

---

<sup>13</sup> Digital Imaging and Communications in Medicine (standard for the communication and management of medical imaging information and related data).

1356 including a computed and derived quantification of blood flow to aid the clinician in the  
1357 assessment of coronary artery disease.

1358 Based on the above definition statement the SaMD *drives clinical management* for in a  
1359 critical *situation or condition*.

1360 This is an example of a Category III.i SaMD used for non-diagnostic purposes.

#### 1361 Clinical Evaluation

1362 As a Category III.i SaMD it is recommended that the manufacturer perform a clinical  
1363 evaluation providing evidence for the scientific validity and analytical validity of the SaMD.

- 1364 • Evidence of scientific validity may be found in literature searches and clinical  
1365 research that shows that fractional flow reserve (FFR) has been validated through a  
1366 number of clinical studies as a safe and effective means for measuring the extent of  
1367 ischemia in the coronary arteries.
- 1368 • Evidence of the analytical validity may include thoroughly checking that the results  
1369 from multiple executions of the SaMD processing the input and output satisfy the  
1370 expected or desirable properties derived from the software specification or user  
1371 expectations:
  - 1372 ○ Testing demonstrated the appropriate functionality of the SaMD and the basis  
1373 of the computational methods;
  - 1374 ○ Evidence demonstrated the functionality and accuracy of the SaMD output  
1375 compared to ground truth data sets of specific modules and components such  
1376 as automatic and semi-automatic image analysis and segmentation tools;
  - 1377 ○ Testing demonstrated the reproducibility of the SaMD output using CT scans  
1378 from various image acquisition systems by the SaMD;
  - 1379 ○ Quantitative evidence demonstrated the validity of the computational  
1380 modeling measurement methods of the SaMD by comparing the  
1381 computational flow velocity solutions to Laser Doppler Anemometry and  
1382 phase-contrast Magnetic Resonance Imaging (MRI) flow data in an in vitro  
1383 model under steady-state and pulsatile flow conditions.
  - 1384 ○ Evidence of clinical performance was generated by conducting a prospective,  
1385 international, multicenter study. Evidence generated from the study  
1386 demonstrated that the diagnostic accuracy of the lower boundary of the one-  
1387 sided 95% confidence interval exceeds 70%.

1388 8.5 Summary of SaMD Clinical Evaluation recommendation

Legend:							
<p><b>Non-D<sub>x</sub>-SaMD</b> = Treat / Non-Diagnostic SaMD</p> <p><b>D<sub>x</sub>-SaMD</b> = Diagnostic SaMD</p> <p><b>AV + SV</b> = Analytical validity + Scientific Validity</p> <p><b>AV + SV + CP</b> = Analytical validity + Scientific Validity + Clinical Performance</p>			<ul style="list-style-type: none"> <li>Treat:</li> <li>Provide therapy to a human body using other means;</li> <li>Diagnose;</li> <li>Detect;</li> <li>Screen;</li> <li>Prevent;</li> <li>Mitigate;</li> <li>Lead to an immediate or near term action.</li> </ul>	<ul style="list-style-type: none"> <li>Aid in treatment:</li> <li>Provide enhanced support to safe and effective use of medicinal products;</li> <li>Aid in diagnosis:</li> <li>Help predict risk of a disease or condition;</li> <li>Aid to making a definitive diagnosis;</li> <li>Triage early signs of a disease or condition;</li> <li>Identify early signs of a disease or condition.</li> </ul>	<ul style="list-style-type: none"> <li>Inform of options for treatment;</li> <li>Inform of options for diagnosis;</li> <li>Inform of options for prevention;</li> <li>Aggregate relevant clinical information;</li> <li>Will not trigger an immediate or near term action.</li> </ul>		
			<b>Treat or Diagnose</b>	<b>Drive Clinical Management</b>	<b>Inform Clinical Management</b>		
<ul style="list-style-type: none"> <li>Life-threatening;</li> <li>Fragile</li> </ul>	<ul style="list-style-type: none"> <li>Requires major therapeutic interventions;</li> <li>Sometimes time critical</li> <li>Vital to: avoiding death; serious deterioration of health; mitigating public health situations or conditions</li> </ul>	<ul style="list-style-type: none"> <li>Specialized trained users</li> </ul>	<b>Critical</b>	<b>TYPE IV.i</b>	<b>TYPE III.i</b>	<b>TYPE II.i</b>	
<ul style="list-style-type: none"> <li>Moderate in progression</li> <li>Often curable;</li> <li>Not fragile;</li> </ul>	<ul style="list-style-type: none"> <li>Does not require major therapeutic interventions</li> <li>Not expected to be time critical</li> <li>Vital to avoiding unnecessary interventions</li> </ul>	<ul style="list-style-type: none"> <li>Either specialized trained users or lay users.</li> </ul>		<b>Serious</b>	<b>TYPE III.ii</b>	<b>TYPE II.ii</b>	<b>TYPE I.ii</b>
<ul style="list-style-type: none"> <li>Slow with predictable progression of disease state</li> <li>Minor chronic illnesses or states</li> <li>May not be curable;</li> <li>Individuals who may not always be patients</li> <li>Can be managed effectively</li> </ul>		<ul style="list-style-type: none"> <li>Either specialized trained users or lay users</li> </ul>			<b>Non-Serious</b>	<b>TYPE II.iii</b>	<b>TYPE I.iii</b>
<b>Disease Type /Patient Condition</b>	<b>Intervention Type</b>	<b>User Type</b>	<p><b>Document AV, SV and CP -- Independent Review not important</b>                      {For Novel SaMD – Build SV and CP evidence using “Real World” experience}</p>				

1389

1390 **8.6 Glossary of Terms Interpreted for SaMD from GHTF Documents**

<b>Accuracy</b>	The degree of closeness of measurements of a quantity to that quantity's true value. When the output of the SaMD and true value are binary, accuracy is the proportion of true results (both true positives and true negatives) among the total number of output values examined.
<b>Precision</b>	The degree to which repeated measurements under unchanged conditions show the same results (related to reproducibility and repeatability).
<b>Limit of detection</b>	The ability of the SaMD to discern between information-bearing patterns of a clinical condition and random patterns that distract from the information.
<b>Linearity or associated transfer function</b>	The behavior of the output across the range of input data that is allowed by the SaMD.
<b>Analytical sensitivity</b>	The degree to which the SaMD's output is affected by parameters affecting input data including perturbation, image resolution, illuminations, data spatial distribution, data amount, etc.
<b>Sensitivity</b>	The ability of the SaMD to correctly identify across a range of available measurements patients with the intended clinical disease or condition (also called true positive rate).
<b>Specificity</b>	The ability of a SaMD to correctly identify across a range of available measurements patients that do not have the intended disease or condition (also called true negative rate).
<b>ROC curve</b>	A graphical plot that shows the tradeoff between sensitivity and specificity as the decision threshold that separates SaMD's negatives and positives is varied.
<b>Positive predictive value</b>	The likelihood of the patient having a disease or condition given that the SaMD's output is positive.
<b>Negative predictive value</b>	The likelihood of the patient NOT having a disease or condition given that the SaMD's output is negative.
<b>Likelihood ratio</b>	The likelihood that a given results would be expected in a patient with the target condition compared to the likelihood that the same results would be expected in an individual without that condition.
<b>Cut-off thresholds or indices or scales</b>	Cut-off values in relation to the clinical condition and on PPV, NPV and likelihood ratio. These should be established prior to validation and must be justified as to how they were determined and clinically validated.
<b>True positive</b>	A SaMD output which correctly indicates that a particular condition or attribute is present.
<b>True negative</b>	A SaMD output which correctly indicates that a particular condition or attribute is absent.
<b>False positive</b>	A SaMD output which incorrectly indicates that a particular condition or attribute is present.
<b>False negative</b>	A SaMD output which incorrectly indicates that a particular condition or attribute is absent.

1391