



# One Bad Apple Can Spoil A Barrel: Are We Effectively Evaluating Software As A Medical Device (SaMD)?

Luc Curtis-Gretton<sup>1</sup>, Robert Malcolm<sup>1</sup>

<sup>1</sup> York Health Economics Consortium, University of York, York, YO10 5NQ

## INTRODUCTION

National policies and agendas have encouraged the digitisation of healthcare systems over the past decade, a trajectory that is set to continue over the coming years. As such, there has been an increased uptake in the use of software as a medical device (SaMD), which is software intended to be used for one or more medical purposes without being part of a hardware device.

There are many differences between medical devices, pharmaceuticals and SaMDs. These include that SaMDs often target multiple indications, are regularly updated, and are more likely to have non-quantifiable benefits. Therefore, the approach to evaluating and integrating SaMDs should be different to that of other health technologies. The objective of this research was to identify problems and potential solutions when evaluating SaMDs in England.

## METHODS

This research took the perspective of the HTA process in England. We conducted a pragmatic review of publicly available grey literature such as NICE guidelines and processes, government schemes, funding mechanisms, and other published reports and opinion pieces. This was used to summarise how SaMDs are currently being evaluated and the difficulties faced in evaluation, both technically and relating to the internal processes. We also conducted a round of focus groups and interviews with health economic and evidence generation experts. We concluded by making recommendations to improve the evaluation of these technologies.

## RESULTS

Table 1 summarises the key issues facing the evaluation of SaMDs to understand the true impact on the healthcare system. A key barrier is engaging SaMD companies to make sure that the evidence they collect is appropriate to inform decision making. Equally, it is important to avoid the assumption that similar SaMDs are as effective as one another. Population, costs, and health outcomes are key considerations when conducting any evaluation of SaMDs, which differ to standard pharmaceutical and medical device evaluations. Possible solutions to these challenges are illustrated in Figure 1.

**Table 1: Key considerations for economic evaluation of SaMDs**

Feature	Consideration
<b>Population</b>	<ul style="list-style-type: none"> <li>The nature of the intervention may impact access. As a result, the SaMD intervention population may differ to the comparator population.</li> <li>SaMDs can often be used across an array of disease areas or populations. Evaluations may have to implement more pragmatic approaches or focus on populations where the impact is expected to be highest.</li> </ul>
<b>Efficacy</b>	<ul style="list-style-type: none"> <li>High quality trial data may not be available due to issues with blinding and randomisation; therefore, SaMDs may be more suited to real-world evidence studies.</li> <li>Iterative versions of the software may mean clinical data quickly becomes out of date.</li> <li>Companies of SaMDs do not always collect appropriate evidence, or do not understand the evidence they should be collecting.</li> </ul>
<b>Costs</b>	<ul style="list-style-type: none"> <li>SaMDs tend to have a wide range of procurement and pricing structures. These need to be accounted for in any evaluation, including aspects such as license costs, company implementation costs, staff time (to implement and maintain), and any other relevant costs.</li> <li>Societal impact should be included for SaMDs that may have significant impact on public health.</li> </ul>
<b>Health outcomes</b>	<ul style="list-style-type: none"> <li>QALYs or similar health state measurements may not be appropriate for SaMDs; cost-comparison or cost-benefit approaches may be more suitable.</li> </ul>

**Abbreviations:** SaMD, software as a medical device; QALY, quality-adjusted life year

**Figure 1: The three-step solution**



## EARLY VALUE ASSESSMENTS (EVAs): A STEP IN THE RIGHT DIRECTION?

**Case study: EVA guidance on artificial intelligence (AI)-derived software to analyse chest X-rays for suspected lung cancer in primary care referrals [2].**

Software with AI-derived algorithms can be used to automatically detect lung abnormalities in chest X-ray images. The EVA analysed in this case study evaluated 14 different software programmes. Table 2 outlines the differences between two of these software programmes.

The committee recommended that more research is needed on all 14 software programmes and did not appear to differentiate between the recommendations. 'Picking winners' will be a big step towards optimising the use of SaMDs in healthcare.

**Table 2: Key differences in two SaMD technologies**

Software	Differences
AL-Rad Companion Chest X-ray	<ul style="list-style-type: none"> <li>This software is a class 2a medical device meaning it is only installed within the body for between 60 minutes and 30 days.</li> <li>It is a computer-aided diagnosis software.</li> </ul>
Annalise CXR	<ul style="list-style-type: none"> <li>This software is a class 2b medical device meaning it is higher risk because they are usually installed within the body for periods of 30 days or longer.</li> <li>It is a computer-aided detection system.</li> </ul>

## CONCLUSIONS

Previous processes at NICE could not keep pace with new technologies requiring evaluation. EVAs offer a route to generate evidence for technologies where there is highest unmet need. However, a backlog of technologies needing evaluation will develop as more evidence is generated through these processes.

Although progress has been made, there is clear incentive to improve the way in which SaMDs are evaluated, both from a technical and process point of view. We recommend the implementation of more pragmatic approaches to evaluation, informed by the decision problem faced. We also recommend a three-step approach to engage, advise, and support companies and decision makers, with the goal of making more informed decisions on SaMDs.

## REFERENCES

- Horgan D et al. Digitalisation and COVID-19: the perfect storm. *Biomedicine Hub*. 2020 Sep 17;5(3):1-23.
- NICE, Artificial intelligence derived software to analysis chest X-rays for suspected lung cancer in primary care referrals: early value assessment [online]

## CONTACT US

luc.curtis-gretton@york.ac.uk

+44 1904 328028

York Health Economics Consortium

www.yhec.co.uk

Providing Consultancy & Research in Health Economics



**INVESTORS IN PEOPLE**  
We invest in people Gold

