Does this thing really work? What evidence is, and how to use it when working with digital health and welfare technologies.

By the PREVIVE health and welfare technology research group at Mälardalen University

The word **evidence** is not always well-received in health and social care services, even if they are usually familiar with it. Many believe that evidence only comes from controlled, scientific studies that care administrators and practitioners don't have the time, energy, or resources to conduct or interpret. Some may even feel that such studies are less applicable "in real life" due to their rigid methods. Others may be uncertain about the tools or methods being used in their work, but "feel" that they are having a positive effect and therefore avoid digging deeper for more objective measures. Things become more complicated when talking about evidence in terms of digital health and welfare technologies, because while it might be familiar to health care settings, it is not a term that is readily used by many civil servants, or even by researchers and developers in related fields such as IT- and computer sciences.

This short article is an attempt to allay such concerns and lower the bar for creating and using evidence – both as a term and as a basis for decision making – when working with digital health and welfare technologies. We'll use a process to do this that can be described by the acronym **ASAP**:

- **A**sk the right question,
- **S**ource the evidence,
- **A**ssess the strength of the evidence, and
- **P**ut the evidence to work.

(A)sk the right question

Simply put, **evidence is any outcome that supports a theory, claim, hypothesis, or assumption**. With health and welfare technologies, we are often looking for evidence that a tool or method is "effective" or "works well" or is "ineffective", "has not made any difference", or perhaps "is harmful". Most importantly, we must first have a well-defined question or problem for which evidence can be obtained. This requires clearly specifying a few key parameters in our question or problem. We can use the acronym **PICOS** to help us in this task. The letters stand for the following:

- **(P)**opulation: the individual(s) or group(s) of interest. It could be patients, users, workers, professions, or even organisations. The more specific we can define the population, the better: gender(s), age range(s), role(s), organisational unit(s) or other identifiable characteristics can be used to help define it.

- **(I)**ntervention(s): the tool(s), method(s), service(s), or influence(s) we are interested in that may affect (or are intended to affect) the population in some way. An example might be a digital surveillance system that is meant to prevent fall injuries at night in the home.

- **(C)**omparison(s): what we can compare the intervention to. This could be the absence of the intervention (including prior to its implementation), our "service as usual", or another, different kind of intervention. If we use the digital surveillance system for fall prevention example, we might compare this to once-a-night, on-site visits. Without this comparison, we have nothing to relate our outcomes to, and it will be difficult to attribute any potential effects.

- **(O)**utcome(s): the effects or events we are genuinely interested in following in our chosen population – and which could potentially be affected by an intervention. It could be the population's health, well-being, workload, or cost-savings. In the case of the digital surveillance system for fall prevention, we might be interested in the number of fall-related accidents that occur, the users' and/or their relatives' sense of security and safety, the costs for hospitalization because of fall injuries, the resources required for on-site visits, etc. Of course, if we don't measure any outcome at all, then we are truly flying blind.

- **(S)**etting(s): The place or environment of interest where the other PICO variables are residing, taking place, or being measured. Our digital surveillance system for fall prevention was for use in the home, which would thus be a setting, but even the workplace of the organisation managing the system could also be considered a setting.

By combining these parameters in our question or problem, we can then identify what evidence is needed to answer or solve it. A question regarding our example of the fall prevention system might be formulated like this:

Does a digital surveillance system (intervention) reduce the number of fall injuries and reduce hospitalization costs (outcomes) in elderly persons (population) living at home (setting), compared to on-site visits (comparison)?

Practice:

Using an actual or potential digital health or welfare technology in your own organisation, work, or situation as an example, create a well-defined question using the PICOS acronym. Discuss it with your colleagues, friends or family and see if they have any different perspectives on what might be included in the question. Make some small changes to one or more of the PICOS parameters and see if it makes the other ones less or more applicable.

(S)ource the evidence

With a well-defined question, we can then look for evidence. Many things can constitute evidence, but in health and welfare services we generally categorize these into four different types:

1) Experimental evidence: this comes from any studies of a question or problem that use scientific methodology to test the effects of something, often (but not necessarily) conducted by experienced researchers. This evidence is often found in published or unpublished research papers, but also in well-conducted internal R&D projects and reports, follow-up, and evaluation results.
2) Observational evidence: this comes from systematically obtained data that is not based on "testing" or altering something, but through careful analysis might be used to explain something. Many organisations accumulate such data but don't ask the right questions, or conduct the right analyses, to turn it into evidence.
3) Professional expertise: the recommendations and advice given by trained persons that are knowledgeable about the population(s) or intervention(s) of interest, or areas closely related to them such as administrative processes, policy, and decision-making etc. This evidence can

be obtained through structured surveys, interviews or panel discussions, statistics regarding usage or choice etc.

4) <u>User opinions and preferences</u>: systematically obtained from those that may be affected by the intervention. These can be highly individual and do not have to be "knowledgeable" or require specific training to be considered relevant. Surveys, interviews, ratings, or even more unstructured formats like journal entries might be used to gather such evidence.

Other kinds of evidence specific to health and welfare technologies also exists. This might be evidence related to safety or technical stability. Meeting certain certifications (such as CE- or EU Medical Devices Regulation), standards (ISO- or otherwise), or even operational stability can constitute evidence for effectiveness. After all, if a technical system is constantly suffering from bugs, downtime, or other related lapses in performance, then its effectiveness will almost certainly suffer as well. It is important to note, however, that achieving a certification is not a permanent guarantee of effectiveness, safety, or anything else. Most digital technologies interact with other digital or non-digital systems and changing one can affect the other. Continual auditing, follow-up, and assessment must be conducted to ensure that such evidence still exists – something which the EU Medical Devices Regulation now requires from approved operating devices.

Practice:

Again, using an actual or potential digital health or welfare technology in your own organisation, work, or situation as an example, identify potential sources of experimental, observational, professional, user-based, or other kind of evidence that might help guide decisions. Where can the evidence be found? Where could it be generated? Where could it be used? Who could help in making it available and understandable?

(A)ssess the strength of the evidence

Evidence should also be defined in terms of its strength. Strong evidence generally has the following characteristics:

1) It is based on outcomes that are **valid**, **reliable,** and **responsive** to change. **Validity** means we are measuring the right thing. A valid measure of our fall prevention system's effectiveness would be the number of falls that occur during its use. Outcomes are **reliable** if we can repeatedly measure them in the same manner several times under stable conditions, and the result is similar or the same. Our body mass, as measured by an accurate scale, should have high reliability if we step on it several times in a row. And an outcome is **responsive** if it is sensitive to changes that may affect it over time. For example, our fall prevention system would be responsive if it could detect small changes in a person's gait when getting out of bed, compared to the last time they got up.

2) It is based on several sources. Conducting experimental studies – where an intervention is conducted in a controlled manner and the effects measured - provides strong evidence. A systematic review or meta-analysis, which gathers several such studies and compares their results, therefore provides some of the strongest evidence. Adding internal reports from organisations and other so-called "grey literature" sources can further strengthen the evidence. Systematically gathered observations are less strong, but still constitute useful evidence. So does gathering expert opinions and advice. Personal experience, judgement, and

intuition are, on their own, neither outcomes nor scientific knowledge, but if they are defined as important in certain outcomes and measured, collated, and interpreted systematically, they constitute valuable evidence as well. If a number of these different sources, when put together, all seem to point in the same direction, then evidence is strengthened considerably.

3) It is gathered from a setting that is like the one where the evidence is going to be used. We may have gathered some very convincing evidence from one country, but if that country's health care system, technical infrastructure, or even training of personnel is vastly different than our own, then the evidence may not be as strong anymore. The same kind of evidence gathered rigorously within our own setting can therefore be considered much stronger. This is known as the **transferability** of evidence – whether it can reasonably be applied across settings and the same effects expected.

Whether we have obtained a lot or a little evidence, we still need to arrive at an agreement about its' usefulness and strength. This is best achieved by allowing all important stakeholders with an interest in the result to assess it. **Evidence is not fact; it is simply a majority consensus at that time and place.** The group of stakeholders doing the assessment agree that a particular outcome supports that something works or does not work. If a new outcome shows something different, then our consensus can change or disappear. It also means that even if we agree that something works here and now, we cannot guarantee that the same consensus will exist *among those people over there in a few months*. They will need to see similar outcomes – the evidence – and agree that it supports that something does or does not work.

So, evidence is fluid. It can get more and less certain, and more or less supportive of our statements, as the outcomes we are interested in change. A greater number or strength of outcomes pointing in the same direction generally gathers more consensus, and the evidence is viewed as stronger. Uncertainty or discord in outcomes means uncertainty. Some may hold fast to their interpretation and say one thing, while others will interpret it differently and say another. Consensus is not achieved, and the evidence is viewed as weak.

Consider:

Use one or more of the outcomes you stated in your well-defined question task. How would you measure them? Would they be valid, reliable, and/or responsive to change? If not, what other outcomes or measurement methods might be more appropriate?

Then, consider the sources of evidence you identified in the next practice task. Does any of it come from your own situation or setting? How would you present it to achieve consensus about a digital health or welfare technology's effects? What is needed to get everyone "on board"?

(P)ut the evidence to work

When we can establish this understanding of the evidence, then we can use it to work "evidence-based". This means both using, creating, and updating evidence continually to support decisions regarding the use of tools or methods.

We can start by using evidence generated by others in making decisions. This might be when we are thinking about procuring a new technology or assessing several technologies to choose the best one. This can be achieved in the following steps:

1. We identify our intention for a specific technology. Why do we need it? What do we want it to achieve?
2. We identify the stakeholders that have a likely interest in the effects of the technology – decision-makers, workers, users, patients, and/or others, and invite them to provide input in the coming stages.
3. Based on our intention, we formulate our problem or question with the help of the PICOS acronym. Most important here is that we clearly identify our outcomes of interest for the technology – to find out if it is having the effect we intended it to – and make sure the measurement of these outcomes is valid, reliable, and sensitive.
4. We go looking for existing evidence, from the various sources described. If needed, we bring in expertise from within or outside of the organisation to help gather and assess it.
5. We compile and weigh the strength of the evidence and present our findings to the stakeholders for their opinions. We describe the level of consensus that results and complement or re-assess the evidence if other questions or concerns arise within the group.
6. We use the results of this process to make our decisions about the technology.

Reflect:

Consider a digital health or welfare technology that you are already familiar with or using. If you had gone through the process above, do you think it would be accepted and used in the same way? If not, what would have been the key step that resulted in the change? Does that step take place in your setting, work, or situation?

Think about when the above process could occur slowly, and when it could occur quickly. What would be an example of where the process could be completed with relatively little effort, or where it would require a great deal of coordination and resources. What would be the most work-intensive step(s)? Would it be worth the effort, in your opinion?

If there is a lack of evidence, or if even if there is lots of evidence from other places, we also want to explore the possibility of *generating* our own evidence. A key principle of what is called *evidence-based practice* is that practice should be based on the most up-to-date and trustworthy scientific knowledge, but the word *scientific* does not necessarily mean it is knowledge created by researchers at an academic institution. While it may help to consult researchers as they are trained in rigorous methods, it is not a necessity. Strong evidence can just as easily be created by a team of social workers, nurses, or administrators. This can be achieved in the following steps:

1. We systematically define our outcomes of interest when using the technology, with all the necessary stakeholders participating. The earlier this is done the better, although it can be done after a technology has been implemented if necessary.
2. We identify the methods for measuring these outcomes, and that they are valid, reliable, and responsive to change. We might obtain measurements via administrative data, external

statistics, user or worker surveys or interviews, or from data generated by the technology itself.

3. We establish a plan for measurement: how often, how much, and for how long. If we are interested in changes in health or cost effectiveness, for example, then we might need to measure for longer periods of time than other outcomes. If we are using a health or welfare technology that is not yet in use, we can measure these outcomes before it is implemented, to have a "control" period that we can compare the technology's effects to once it is operating. If the technology is already implemented, we can see if measurement data is saved somewhere prior to its implementation.

4. We establish a plan for analysing our measurements: what is considered an important or "clinically significant" change? Is it a user moving from one category to another? A percentage reduction? A critical or threshold level? We determine how often the analysis should be conducted, and by who.

5. We communicate our findings from the analyses – this is now evidence. Who should see the results, and when? *Pro memoria* or real-time data visualisations might be effective for getting the word out internally, but are there other groups or stakeholders that might benefit?

6. We use the results of this process to make decisions about continued or altered use of the technology.

Get the knowledge "out there"

The evidence we have generated doesn't have be published as a scientific study to be considered as such, either. The benefits of publishing our findings are that we are letting our outcomes, and the methods for obtaining and interpreting them, be reviewed by others. If they also agree that our methods and interpretations are correct, then we have increased our consensus in a very impartial and transparent manner, and the evidence gathers strength. By publishing this, we then allow even more people to assess our methods and interpretations, and even to repeat them and see if they arrive at the same result. Evidence becomes even stronger as more arrive at the same result, and in turn our basis for decision making as well. So don't hide results – getting them "out there" is truly putting evidence to work! Findings can be published on an organization's external website, as a conference presentation, or in a branch-specific or regional publication. Such reports are often up-to-date, highly context-specific, and thus relevant for many other practitioners and decision-makers. We are contributing to the development of evidence-based knowledge!

A final word

Hopefully this short article will have shown that working in an evidence-based manner is not restrictive or rigid. It allows us greater freedom to investigate, explore, and change our work with less frustration, uncertainty, and resistance. It promotes agreement and decision making in a structured, inclusive, and transparent way. It develops our competencies and knowledge of our work, no matter at what level.

Most importantly, **evidence allows us to focus our energy on that which yields the most benefits.**