Measuring Impact

A toolkit for your digital health community

UNIVERSITY OF LEEDS

Shift.ms

Foreword

Shift.ms is a social network for people with multiple sclerosis (MS) founded in 2009 by MSers, for MSers. The charity supports many thousands of recently diagnosed people across the world as they make sense of MS.

This toolkit lays out a set of principles that can be applied to evaluate digital technologies often used by health communities. It is also a summary of Shift.ms' learning from evaluation projects either self-initiated or conducted in partnership with academics and external evaluators.

"When diagnosed with multiple sclerosis in my early twenties, one of the first things I did was to look for other people in a similar situation who could help me make sense of what was going on. After searching my local area and looking online, I struggled to find people living with MS who were my age and facing similar challenges. To fill this gap, I decided to set up my own digital community and founded Shift.ms.

Back in 2009 I couldn't have imagined the scale to which the small online forum, launched with the support of a group of friends and family could have grown. The community is now over 40,000 strong with members all around the world. We hope that through the development of this toolkit and partnership with the University of Leeds, some of the learnings gleaned over the last ten years can be adopted by other grass-roots or more established digital communities."

George Pepper, founder and CEO, Shift.ms

"Digital health communities have become fundamental in our society. They are often run by a bunch of people with limited resources but reaching thousands around the globe through their immediate, independent person-centred peer-support. Academics want to work collaboratively with global digital health communities so they can better plan, manage, conduct and use evaluation to maximise and demonstrate their impact."

Dr Ana Manzano, Researcher, University of Leeds.

Funding and licence to use the toolkit

Funding: This toolkit was funded by an ESRC IMPACT Acceleration Account Grant at the Leeds Social Sciences Institute. The toolkit was co-produced with Shift.ms and was also supported by the School of Sociology & Social Policy and the Leeds Institute of Health Sciences at the University of Leeds (United Kingdom). The views expressed are those of the authors and not necessarily those of the funder.

As part of this licence, we ask you to cite this toolkit as: Manzano A, Hentschel J, Varela A, Webb EJD, Shift.ms. 2021. Measuring Impact: A Toolkit for Your Digital Health Community. University of Leeds and Shift.ms.

Disclaimer: Every effort was made to provide accurate information at the date of publication (March 2021), however, some of the facts and sources presented here may change over time. In this fast-moving digital world, health communities keep adapting and innovating and many other challenges and evaluation tools will appear in the future, which are not discussed in this toolkit.

Introduction:
Who is this toolkit for

This toolkit is for people who work at, volunteer with and/or run digital health communities. Digital health communities (also known as online health communities) are "virtual platforms that enable members to interactively discuss health-related matters in order to get knowledge or psychological support[1]." These can include a breadth of organisations, from small grassroots to bigger, established organisations based in many different countries.

Digital health communities are diverse, some are dedicated and run only by patients and users; and others aim to improve communication between specific healthcare professionals (e.g. nurses, physicians). They may be moderated (e.g. by providing and monitoring rules) or unmoderated, and may be run by professional facilitators, untrained or trained peers. In this toolkit, we focus on patient peer-to-peer communities offering support mainly through digital means.

Despite their undeniable impact on society, fast-moving digital health communities often struggle to demonstrate how they impact people's lives and their relationship with their illness. This toolkit aims to fill this gap and provides a set of principles and resources to support evaluation activities, supported by real case studies.

How to use this toolkit?

This toolkit is designed to be read in order. Each section is complemented by Shift.ms case studies, tips and further sources that our team found particularly helpful as they worked through their own projects. At the end of the toolkit, you will find a "Self-Evaluation Checklist" summarising the core of the toolkit in 15 steps that make up the evaluation process.

What do we mean by impact in digital health communities?

Since digital health communities are diverse, their aims are also varied, and their impact is related to those aims. In general, digital health communities aim to offer social support that can help people gain power to deal with their health and with healthcare services. This in turn can have a positive impact on people's health.

The support offered and gained through peers in digital platforms consists of sharing first-hand knowledge and experiences, providing information about living with the condition day-to-day and disease course (diagnosis, treatments, prognosis).

In practice, this support can look like tips on practical help, social interaction through online and offline conversations, and emotional support. All these can help people cope with new and difficult experiences for their everyday lives and improve their health and wellbeing.

What type of impact can digital health communities have? Shift.ms case study

People affected by MS often feel isolated as a result of the condition, making the reduction of social isolation and loneliness a key impact area for Shift.ms. As well as being accessible at any time, Shift.ms aims to offer highly personalised levels of support, with members able to access others who relate to their specific demographic, lifestyle choices and health challenges. Shift.ms offers an online forum, which is a place to ask questions and share experiences and information. They provide an 'always on' service that can dovetail with more traditional health services. It also aims to increase confidence, knowledge and strategies to better manage and live with health challenges. Knowledge and lifestyle tips picked up through experience are commonly posted, while professional, scientific and health knowledge is often collectively digested and linked in different formats including short videos and interviews with community members, healthcare professionals and scientists.

Consider the following questions:

- Why do you want to evaluate now?
- What do you want to learn?
- What priority areas should you focus on?
- How will the results of your evaluation be used?

How to evaluate digital health communities' impact?

For digital health communities, "evaluation" is the process of measuring and monitoring the impact of the services they offer through virtual platforms against their community objectives. This can support decisions about improving their services.

Evaluation can be done under many different labels, such as: review, assessment, impact analysis, social impact analysis, appreciative inquiry, and cost-benefit assessment[2].

Evaluation activities can be done before, during and after implementation of new services, and regular evaluation should be a long-term objective of digital health communities, rather than being seen as one-off events.

Digital health communities may wonder whether they should commission an external evaluation or do their evaluation themselves. Even if some evaluations are outsourced to other organisations, digital health communities should always aim to do some self-evaluation themselves.

This guide to evaluation written by the Big Lottery Fund includes information on what evaluation is all about and ways to plan evaluation from the start: https://evaluationsupportscotland.org.uk/resources/a-guide-to-evaluation-by-the-biglottery-fund/

This collection of videos addressed at third sector organisations reflects on pros and cons of commissioning evaluation vs carrying out self-evaluation: https://evaluationsupportscotland.org.uk/resources/1-of-4-should-i-self-evaluate-orcommission-an-external-evaluation/

How to get started? Step back!

First, take a step back and consider the bigger picture of the digital health community. This is done by describing in detail the aims and services provided by the digital health community that is going to be evaluated. Different people may have different views on these and it is best to do this in a group and to start with small areas of the service.

If possible, assemble an evaluation support team involving roles such as:

- Digital health community decision-makers (CEO, CTO, directors)
- Service designers
- IT managers
- Stakeholder engagement team such as fundraising or operations
- User representative of the digital health community

Logic maps (also called logic models or theory of change) can be useful tools to step back, since they provide an overview and help identify priority areas for evaluation. This is done by visually identifying how digital health communities' activities or services (forums, blogs, videos, etc.) aim to contribute to community objectives, and how the results obtained after conducting those activities eventually produce impacts and how they look.

Logic maps are drawn as simple diagrams to spell out those chains of objectives, resources and activities that are expected to produce short-term, medium-term and long-term intended results (outcomes) and overall impact.

A comprehensive guide for logic mapping addressed at policy makers, evaluators, local authorities and partnership organisations can be found here: https://www.gov.uk/government/publications/logic-maping-hints-and-tipsguide

A comprehensive guidance for logic Check out this guide on how to put together a basic logic model to help your organisation think about your objectives, outcomes and activities: https://evaluationsupportscotland.org.uk/resources/ess-support-guide-1c-developinga-logic-model/

Goal	Objectives	Resources	Activities	Short-term outcomes	Long-term outcome	Impact
To increase the health of the population.	To promote the benefits of carrots. To increase carrot consumption in children and young people.	Named staff to work on carrot activities 20% of the time. IT and media support. 5% of annual budget.	Media campaign, blogs by carrot supporters. Videos with health professionals. Tab in main website.	Increase mothers knowledge about carrot benefits in children, increase carrot consumption in target population by 10%.	Improve population nutritional status.	Improved health.

How to narrow down evaluation objectives?

The next step is to decide and agree on the purposes or objectives of the evaluation, which are linked to the mission (aim, that is, what you hope to achieve) and objectives of the digital health community (that is, how you are going to achieve it). **Typical objectives** of evaluations in digital health communities are:

- To inform decision-making related to improvement such as starting, changing, stopping or confirming services offered to the community.
- To demonstrate achievements so that lobbying and advocacy can be supported, e.g. for better access to services or treatments.
- To build trust in the digital health community by demonstrating that expectations are met.
- To ensure accountability for internal purposes or to demonstrate it to funders and community members.
- To ensure diverse perspectives are included so all intended beneficiaries are reached.
- To contribute to the generation of new evidence to influence practices and policies decided by others outside the digital health community.

Digital health communities are varied ranging from very small grassroot organisations to larger, established ones, to those working in partnerships. But no matter how small evaluations are, they can be time-consuming and expensive. Often organisations cannot complete their evaluations because of this. Don't be put off and try to:

- **Have a reality check.** Assessing available resources helps narrow down objectives and build commitment from the evaluation team and stakeholders.
- **Be flexible.** Make the evaluation plan adaptable to different circumstances and organisation size.
- **Collaborate.** Think about partnering with external organisations, e.g. universities, other third sector organisations, external evaluators, etc.

The Better Evaluation Rainbow Framework is a good resource to get started. It organises useful processes and methods to start an evaluation: https://www.betterevaluation.org/rainbow_framework

How to narrow down what is going to be evaluated?

It is often unfeasible to evaluate the overall impact of a digital health community, so it makes sense to break it down into smaller evaluation areas. It is useful to express these areas as questions and to then link these questions to specific tools to answer them (measurement tools).

Evaluation questions are tools to fine-tune the focus of the evaluation. They could, for example, focus on how the digital health community works, what outcomes are observed, or how the digital community is reaching diverse users differently.

To write evaluation questions, it is useful to return to the components of the logic map representing the objectives and services offered (see previous section). Some of these components are related to processes and others are related to outcomes.

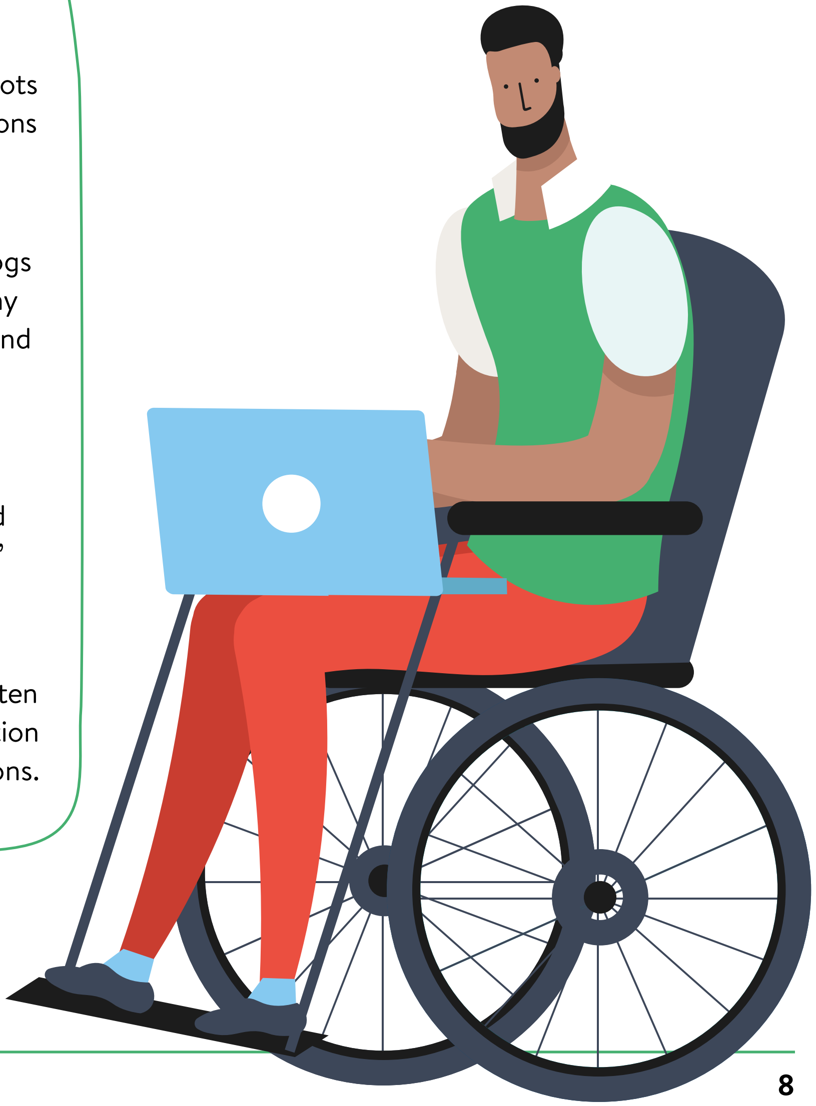
Example:

The digital health community wanting to promote the health benefits of eating carrots could have the following evaluation questions about process or outcomes/impacts:

Process evaluation questions:
To what extent are carrots' videos and blogs delivered in a timely fashion? How and why does the knowledge gained from videos and blogs change community members' food consumption behaviours?

Outcomes evaluation question:
To what extent are the carrots' videos and blogs having a positive impact on families' knowledge about carrots' benefits?

Good evaluation questions are:
Informed by the evaluation objectives. Written with the help of your stakeholders. Evaluation questions are broader than survey questions.

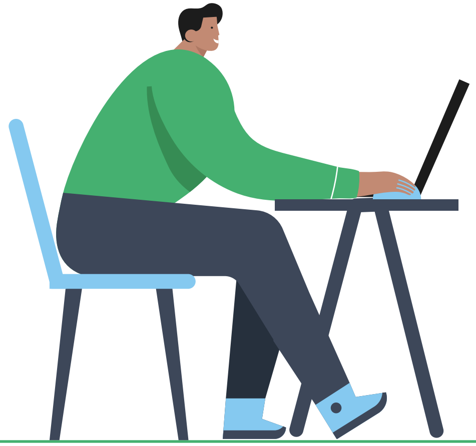


Write the logic map of your intervention first and then relate the logic map components (resources, activities, outcomes, etc.) to possible evaluation questions. Check this table[3] for examples of general evaluation questions typical in digital health communities:

It is best to focus on a small number (maximum 5) of overarching evaluation questions with more detailed questions being linked to each of those.

Useful resources for setting up evaluation questions are: Evaluation Building Blocks: A Guide by the Kinnect Group: http://kinnect.co.nz/evaluation-building-blocks-a-guide/

Better Evaluation: Specify the Key Evaluation Questions https://www.betterevaluation.org/en/rainbow_framework/frame/specify_key_evaluation_questions



Type of Evaluation	General Evaluation Question	Link to Logic Map Component
Process Evaluation	Are human/material resources adequate to implement this new service/idea?	Resources
	Is the service (and activities) implemented as expected/planned? Why?	Activities
	How much was offered? How many were reached? Who was not reached? Why, for whom and in what circumstances?	Objectives
Outcome or impact evaluation	Any change in knowledge, policy environment? What intended and unintended outcomes (positive and negative) were produced? To what extent can changes be attributed to the new activities? What were the particular features of the new activities and context that made a difference? To what extent can changes be attributed to the new activities? What were the particular features of the new activities and context that made a difference?	Short-term outcomes
	Any change in behaviour?	Intermediate outcomes
	Any change in health status?	Long-term outcomes
	Any change in population health status?	Impact

How to choose indicators for measuring digital health communities impact?

Indicators help to decide what success looks like for the digital health community's objectives and activities. To identify whether things have improved or got worse or how they can be improved, it is important to agree on criteria and standards before any data is collected for an evaluation. For example, digital health communities are known to be motivated by four types of social support[4]:

Informational	**Emotional**
Esteem	**Network Support**

Possible **evaluation criteria** or aspects that are important to consider when assessing whether a service/activity has been successful or not, for whom, why and in what ways are:

- **Positive outcomes.** E.g. Should digital health community informational support be judged in terms of success in promoting interactions between community users and healthcare services? Or success in supporting community users to engage in work, educational and leisure activities? If it is both, how should both be weighted?
- **Negative outcomes.** E.g. Producing too much content (e.g. videos, blogs) about a new controversial theme (e.g. a new unlicensed treatment offered illegally in some countries) may produce negative unintended effects (e.g. losing community trust), as well as positive intended effects (informational and emotional support in terms of hope).
- **Costs and benefits** and how they are distributed. E.g. How is the website content targeted so that some community members (e.g. newly diagnosed users, in the case of a digital community of people with long term conditions) receive more benefit?
- **Resources and timing.** E.g. Do evaluation results need to be achieved within a certain timeframe (e.g. before the next funding application round)?
- **Processes.** E.g. Making website content accessible to users with visual or hearing impairments.

Check out these one minute videos on what outcomes and indicators are and how to set them: https://evaluationsupportscotland.org.uk/evaluation/evaluation-pathway/setting-outcomes-indicators/

Evaluation standards are the levels of performance necessary for each of those criteria. For example, for a digital health community aiming to increase network support through a forum, what level of community engagement is needed for it to be considered successful? An increase of posts? How many posts? By how many users, in what geographical regions or age range, in what timeframe?

Criteria and standards need to be agreed before deciding what data will be collected and how. The combination of this information will help form an overall judgement of success or failure.

Examples and guidance on how to decide on standards, evaluative criteria and benchmarks can be found here: https://www.betterevaluation.org/en/evaluation-options/benchmarks_standards

Generating new quantitative data for measuring impact

Outcome and impact data can be newly generated and/or can also consist of data that are routinely generated by website usage. New data can be collected by designing your own surveys or using off-the-shelf validated tools. Both have advantages and disadvantages.

1. Designing questionnaires and surveys

A **questionnaire** is a list of questions and answers aimed at gaining specific information about attitudes, preferences and factual information from a selected number of respondents (called a **sample**). The answers chosen by this smaller population are supposed to represent the answers of the wider population in question. The answers can be unrestricted (free form text), allowing people to express their views and experiences with few restrictions (qualitative) or; restricted to limited options of pre-determined answers (quantitative). A questionnaire may be part of a wider **survey**. A survey is the process of collecting, analysing and interpreting statistical data such as questionnaire data, but it often involves more than one form of data collection.

Designing a questionnaire may seem an easy task but question wording, ordering, exhaustiveness of answer options, sampling decisions, etc. all require advanced technical knowledge. Errors can lead to results which aren't helpful, and can lead you to the wrong conclusions. A high-quality questionnaire requires thorough testing and validation processes, which can be time consuming.

Take into consideration ease of use and people's attention span online. The more concise and easier the questionnaire is to use the better!

Case study. In MS, diagnosis can take a long time and people might be visiting the digital health community to get information about the disease, to exchange symptoms, to meet others who are in similar situations, etc.

If in an evaluation questionnaire, a question asking whether a person has MS only includes "yes" and "no" as possible answers, neither option applies to someone who is waiting for a diagnosis or might suspect they have MS. This would reduce the accuracy and usefulness of the data.

If a question does not offer exhaustive answers, in other words, if the answers do not include an option reflecting the experience of everyone who might take the questionnaire, measurement errors will occur.

Snapshot or longitudinal?

In longitudinal survey studies, data about the same people is collected over time, in regular intervals (e.g. asking the same person “how lonely do you feel at the moment?” every 12 weeks for 12 months). Key decisions are: how many rounds of data collection will be conducted, over what period of time and how will data be linked from earlier or later rounds?

Snapshot studies (also called cross-sectional) collect data from a group of people at one point in time. In retrospective studies, information is collected about their past events and circumstances (e.g. have you felt less lonely since you started using our online forum?).

Longitudinal studies can be robust but slow, expensive and highly complex. Over time, some participants will drop out of the study and datasets are more complex to manage and analyse than those for snapshot surveys.

Do you need baseline data? Baseline data are measurements of key conditions before a project begins, from which change and progress can be assessed. Sometimes baseline data is available, but sometimes a baseline study is needed. However, baseline data is not always necessary, possible or useful. The value of baseline data for either longitudinal or snapshot studies varies depending on the outcomes measured and the type of evaluation. Check out this resource: https://www.betterevaluation.org/en/resources/guide/baseline_basics

2. Using non-commercial and commercial off-the-shelf questionnaires

Pre-existing measuring tools can be commercial or non-commercial. Using these can have major benefits. For example, if you want to measure general health, you might be better using validated standardised tools like EQ-5D[5] or SF-6D[6] rather than reinventing the wheel. There are many validated measures in the public domain which are free to use and have been rigorously validated across many different contexts. If you do not have advanced expertise in questionnaire design, time and resources, you are better off using off-the-shelf scales. You can use them to collect baseline data, for snapshots or to measure outcomes over time.

Caution is needed when purchasing licensed commercial data collection tools, which commonly have a number of problems:

- They are normally easier to buy than to use.
- Advanced quantitative technical expertise is still necessary to analyse the data.
- Before you buy surveys, you often cannot see what they look like (how long they are, type of questions, etc.). Some of these products may have been designed for analogue data collection and never been tested with digital users.
- It is difficult to ascertain suitability prior to purchasing the products; consequently, this will often only be properly ascertained once the survey has already been implemented and data has already been collected. Changing the product at this point is usually expensive, frustrating and even impossible.
- Commercial products are likely to limit your control at implementation (since you cannot change questions) and analysis (calculations and scoring algorithms can be hidden and still controlled by the vendor). The analysis provided by the vendor can be too basic for the digital health communities' evaluation needs.
- Customer support from the product /software vendor can be slow and/or a costly upgrade to a more expensive support package.
- If the company shuts down, all the data collected by you/as part of the evaluation may be lost.

3. Formatting and distribution channels for questionnaires

Online survey tools can facilitate the formatting, distribution, collection and analysis of quantitative data. There are many free tools available and it is worth spending time making sure that you choose the right one for your needs, for instance, one that features mobile-friendly design.

It is essential to check that they are compliant with all relevant data protection and privacy regulations. It is also important to be aware what regulations apply. For example, data collected from EU citizens has to be handled according to the General Data Protection Regulation (EU GDPR), even if you are based outside of the EU.

Some subscription survey tools are designed for the public sector and have free trials. Google Forms is a free online survey tool but there are many others available. Many universities subscribe to www.onlinesurveys.ac.uk/, an online survey tool designed for academic research, education and public sector organisations.

How do you share and distribute online questionnaires?

Launching your online questionnaire so it is open for responses does not automatically distribute it to your target population. An online questionnaire can be distributed in many ways, such as:

- Sending it by email by creating an email invitation that includes the URL to the online survey form, and distributing it to a list of known email addresses. The key advantage here is the ability to target the distribution to specific participants. Ensure you have people's consent to send email invitations, as this is required by GDPR.
- Embedding it into your website. Website traffic is likely to include a relevant population, but this option has fewer targeting features. To increase targeting, questionnaires can be added to specific pages, included as pop-up invitations for first time visitors/registered users, or set up as follow-up pages for users who have participated in website activities (e.g. watching a video, writing their first post).
- Embedding it in blog posts where the details and background of the study can be included to help participants understand the study and also target specific participants.
- Sharing it via social media via a direct link or a partnership with influencers/collaborators.

All of these can be used individually or combined. Visualising distribution summaries to track response metrics is a useful tool to see which channel is most effective. All channels have advantages and disadvantages and they all require different time, budget and human resources. All channels are subject to participation self-selection and therefore, bias (influences on responses)[7].

If digital health communities are trying to reach potential participants who regularly use their website, embedding the evaluation survey into the website (where they already are) rather than sending out an email to everyone (active and inactive members) might be a good idea to improve access to the key population and to optimise response rates.

Collecting data about the impact of digital health communities is an onerous and complex exercise. This requires technical knowledge (e.g. IT, advanced quantitative data analysis skills) and considerable investment in resources. Some communities may already have those skills in their current workforce, but if not, it is recommended that additional funding is specifically dedicated to this. It is unlikely that such a difficult task could be conducted effectively as an add-on to current staff roles. Collecting data may seem easy, but doing it in a way that produces relevant, reliable, high-quality data is difficult and requires many considerations. It might be worth hiring an expert, even just for consultation.

Always run a pilot study or evaluation test-run to identify potential problems and revise the survey (not only questionnaire wording but question order and number). This should include implementation of the survey and analysis of the different data at all time-points to examine linkage issues. Key phases to test run are:

- Participant recruitment. E.g. which distribution channels are more effective in attracting the target population.
- Testing the surveys or impact tools. E.g. exhaustiveness of answers, understanding of questions, time taken to complete questionnaires, willingness to consent.
- Data entry and longitudinal data linkage analysis. E.g. identifying errors such as incorrect, incomplete, irrelevant, duplicated, or improperly formatted data, correcting, modifying or deleting data, or manually processing data to prevent the same errors from happening.

Dealing with data routinely generated by website usage

It is often challenging to work out how to get the data that is important to your organisation into a place that is useful and easy for the team to access, analyse and action.

Customer relationship management (CRM) software is a technology that allows digital health communities to analyse user interaction with their website content. CRM software records customer contact information (e.g. email, website social media profile, personal preferences on communications) and the website (or company) activity. A CRM can present this data in customisable dashboards to give you a better understanding of your service over time.

There are many companies that offer these services for non-profit organisations and they tend to work via a subscription where organisations pick and choose the services they are interested in and pay a monthly/annual fee. One of the challenges of using CRM companies is that their matching methods and algorithms tend to be hidden and controlled by the company. These help determine how a specific field in one record is compared to the same field in another record and whether the fields are considered matches.

Case study. Shift.ms wanted to know how users behave on their website (e.g. how many posts they create, how many messages they write) alongside general demographic information (e.g. age, type of MS, when they were diagnosed).

They decided to subscribe to a CRM software company that collects this data and organises it so that Shift.ms can log onto their platform and access this information. Shift.ms can now look up specific users via the CRM platform or pull reports about specific aspects. For example, how their user demographics changed in the last 3 months, how many people with a certain characteristic posted something in the last week, etc.

These reports can be configured so that they update themselves automatically when additional data becomes available, e.g. when new users join.

Check out these resources discussing features of the key CRMs for charities and non-profits: https://fundraising.co.uk/2017/01/16/top-10-crms-charities-nonprofits/?cmpredirect https://charitydigital.org.uk/topics/topics/thebest-crm-systems-for-charities-7626

For an explanation about the problems with correlation, check this website: https://hbr.org/2015/06/bewarespurious-correlations

Linking new quantitative data with routinely generated website usage data

Data linkage is an important research tool. It can be useful because it could uncover/reveal additional information about users, especially when it is correlated with impact data collected through other tools.

Linking questionnaire results to other routinely collected data by digital health communities is exciting but this can also be highly complex. Collecting data from the same participants at several points in time (also called "longitudinal" data) can be even harder. Small matching errors can result in reduced samples and loss of vital power needed (the probability that your results are unreliable) to draw robust conclusions. Remember: Correlation is not causation!

Case study. Shift.ms' data linkage between a survey about loneliness and wellbeing and website usage data collected with CRM could help answer questions such as: "Do users who are active on the website (creating posts, messaging, liking posts, etc.) have lower loneliness scores?" "Are people with one type of MS more or less likely to have improved wellbeing scores than people who have other types of MS?" It can also be useful for analysing response and attrition rates during the pilot phase "What characteristics make people more likely to take part in the survey?" This could also shape the evaluation after the pilot phase ("Do we need to change our data collection approach to reach the people we are interested in?").

Ethics, content management and user anonymity

Ethics considerations are important issues in all evaluations. Digital health communities deal with big data and many of the ethical issues can tend to focus on these.

Key ethical principle: Data Protection from the start

- Data warehousing and compliance need to be a must from the offset for digital health communities, this is part of the duty of care towards community members.
- General Data Protection Regulation (GDPR) is the legal framework that currently sets guidelines for the collection and processing of personal information in the UK. If the data that is being collected is anonymous and does not contain any personal data, GDPR compliance is not necessary.
- Although GDPR is a complex process, it helps your users trust you with their data.
- Most digital health communities collect personal data (age, region, etc.). Different self-evaluations for digital health communities will have different GDPR challenges.

In the UK, check the Information Commissioner's Office (ICO): https://ico.org.uk/for-organisations/data-protection-self-assessment

A GDPR compliance checklist for data controllers can be accessed here: https://gdpr.eu/checklist/

Key ethical principle: understanding anonymisation and pseudonymisation

Completely anonymised data takes away any means of identifying individuals. Pseudonymisation, instead, takes the most identifying characteristics of the data and replaces them with one or more artificial identifiers, or pseudonyms, for example by replacing a name with a unique number. This means that users are less identifiable, but it is still possible to identify individuals with a code.

Be aware of the **difference between anonymised and pseudonymised data**, and which one describes your data. If there is any way that someone could link one or more pieces of data back to an individual, your data is not anonymised. This includes having enough details available to narrow down a data point as coming from a single individual. E.g. if you have a diagnosis, age, gender and location, there may be only one person in that location with that diagnosis, age and gender. The rarer a disease is, the more difficult it is to anonymise data, since the lower the number of people that have a condition, the less information is needed to narrow it down to an individual.

How **data routinely collected by website builders** (e.g. WordPress) is used for evaluation purposes is protected under relevant laws (e.g. GDPR); and ethical principles to respect people's privacy. Service evaluation, audits and monitoring conducted by digital health community staff follow standard data protection regulations.

If data is shared with external evaluation collaborators (e.g. contractors, health organisations), then you may have to fully anonymise your data. Understanding how your website builder anonymise data is key. For example, if using WordPress:

- WordPress IDs are used to identify users on websites. The IDs and any information connected to that user must only be accessed by people who have access to the back-end of the website, e.g. developer, website administrator.
- WordPress IDs cannot be used by other people outside the digital health community to identify users. This does not necessarily mean that the data collected from these users is fully anonymised. It does refer to individuals and is connected to their data on the digital health community, even if others cannot link it or access it.

You may be able to anonymise your data after linking two or more sources. For example, you could link survey data to routine website usage data collected by a content management system using WordPress IDs (or similar IDs from other companies). Even though you don't know the names of your participants, the data is only pseudonymised as the content management company has a key linking the WordPress ID to the individual. If you use a random number generator to create new IDs, and permanently delete the WordPress IDs from your data, it may be anonymised since there is no way to link the new random IDs to the individual.

Key ethical principle: informed consent and surveys

- When digital health communities consider collecting impact data through questionnaires, they must note that participants must provide consent to allow them to collect and handle their personal data.
- The participants can only give consent if they are told about how their data will be used and the purpose of the questionnaire prior to data collection. Their consent must be unambiguous and explicit.
- For online surveys, the consent checkbox must not be selected by default. Participants also reserve the right to revoke their consent at any point.
- If it is impossible to remove participants' data at some point, for example if you fully anonymise a dataset, they should be made aware at which point this will happen.

If working with external partners (e.g. evaluation consultants, online survey companies, universities) to collect or analyse your data, think about how these principles apply. For example, only anonymised data should be shared with others. It is your responsibility to do your due diligence to ensure your partner and anyone you share data with is GDPR compliant. It is good practice for you to have written data sharing agreements when controllers share personal data.

The ICO website is a useful resource for all things data: https://ico.org.uk/media/for-organisations/documents/2618790/data-sharing-code-thebasics.pdf

The internet is a fast moving beast! Check out the Association of Internet Researchers' Ethics Guidelines since they will update them as things change: http://aoir.org/ethics/

Many countries have evaluation societies that produce ethical and practice guidance:

UK: https://www.evaluation.org.uk/professional development/good-practice-guideline/

North America: https://www.eval.org/p/cm/ld/fid=51

Canada: https://evaluationcanada.ca/ethics

How to analyse data

Data analysis is an important part of an evaluation plan (to answer those evaluation questions) but this is often only loosely pre-defined. Appropriate knowledge and tools are needed to analyse quantitative or qualitative data and to integrate both, if different types are used.

Microsoft Word and Excel, Google Drive Spreadsheets, etc. can also be used for descriptive analysis as well as basic statistical analysis. There is specialist data analysis software specifically for quantitative data such as SPSS and R, or qualitative data such as NVivo.

It may be tempting to give questionnaire respondents the option to answer open-ended questions to capture more complex individual thoughts, views and perceptions. While answers to such questions can be rich and informative, analysing free-text data can be an elaborate and timeconsuming process. An analytical strategy must be decided, which may require quantitative (e.g. quantitative content analysis) or qualitative (e.g. thematic analysis) analysis skills. There are many software solutions tools in the market for coding and analysing textual data. In all cases, the data still needs to be coded and interpreted by people. Also, be aware that free-text responses require more effort from participants, and so may be more off-putting compared to multiple choice responses.

If working with external partners (e.g. evaluation consultants, online survey companies, universities) to collect or analyse your data, think about how these principles apply. It is your responsibility to do your due diligence to ensure your partner and anyone you share data with is GDPR compliant. It is good practice for you to have written data sharing agreements when controllers share personal data.

Summary statistics, correlation, crosstabulation and frequency tables are some useful systems to analyse quantitative data. You can find a summary of the basic analysis systems for quantitative and qualitative data here: https://www.betterevaluation.org/en/rainbow_framework/describe/analyse_data

How to analyse data using common software: https://www.betterevaluation.org/en/blog/analyzing_data_using_common_software

For tutorials on key statistical concepts, check these: http://wise.cgu.edu/
https://www.coursera.org/learn/basic-statistics
https://www.coursera.org/specializations/social-science

Quantitative data analysis software overview: https://stats.idre.ucla.edu

Qualitative data analysis software overview: https://www.surrey.ac.uk/computer-assisted-qualitative-data-analysis/resources/choosing-appropriate-caqdas-package

User-led methods and pioneering creative methods

Since digital health communities tend to be led by users, it is important to recognise and use the expertise of service users by also doing user-led evaluations.

Pioneering new methods that are credible and suitable for your digital health community can have a similar or more positive impact in your organisation than traditional evaluation methods.

Users can be involved in many ways but in most cases:

- **Users can play a role** in constructing the process, standards and criteria for evaluation.
- **Users can undertake the evaluation** themselves with some support, usually applying predetermined processes and criteria and evaluation approaches. Empowerment evaluation is an example of this type of user-led evaluation.

Read more about Empowerment Evaluation here: https://www.betterevaluation.org/en/plan/approach/empowerment_evaluation

In the next page you can find a summary of this toolkit advice in a 15-step self-evaluation check-list comprising of questions and assigned actions. We hope you found this toolkit useful, and good luck with your evaluations.

Finally, you can check out this resource on how to report evaluation results and recommendations: https://www.betterevaluation.org/en/rainbow_framework/report_support_use

Case study. Shift.ms has tried and tested many different methods of evaluation and a key learning has been to view methods of evaluation as being part of the core user experience, no matter the platform, as opposed to something separate.

Research tools are designed to best suit the immediate activity. If a closed video call with a group is being hosted for example, a short evaluation shared on screen could be complemented with an open forum for conversation, with qualitative feedback collected in the ‘chat’ function.

Evaluating services on digital platforms without inbuilt means can be tricky. One solution is to factor in the digital tools that are available. With YouTube videos, for example, reviewing comments and replies, alongside performance metrics, gives some insight into the impact of community-driven content.

It also helps if community members at the heart of the content are helping shape/ are aligned with the objectives, so the content can be tailored to encourage community feedback around key evaluation questions.

Self-evaluation checklist

To support you in your self-evaluation process, we suggest following a 15-step guide[8] which outlines the key areas explained in this toolkit:

	Prompt	Toolkit pages	Actions	Assigned to	Due date
STEP BACK (Before starting designing the evaluation)	1. Have you drawn a logic map of your digital health community objectives, activities, outcomes and impact? Who do you need to involve to draw this map?	5-6			
	2. Is the evaluation clearly relevant to the needs of your digital health community, as identified by any form of study, or other evidence and argument?	7-9			
	3. Have you identified what actors outside your digital health community are needed to support your evaluation project?	6			
	4. Have the primary users of the evaluation been clearly identified? Can they be involved in defining the evaluation? Will they participate in an evaluation process?	21			
	5. Have you allowed time to evaluate within the schedule of all other activities? Funding available for the relevant team and duration? People with the necessary skills.	5 and 7			
	6. Do you need to collect some baseline data before designing this evaluation project? If baseline data exists in the form of survey data or CRM data, is the raw data available? Or just a selection of relevant items?	12			
STEP IN (Designing and doing the evaluation)	7. Are there short-term, medium-term and long-term outcomes and impact clearly identified in the evaluation and are the proposed steps towards achieving these clearly defined?	6-10			
	8. Is it likely that the evaluation objective could not be achieved, within the project lifespan? Why? Do you need to narrow it down?	6-10			

	Prompt	Toolkit pages	Actions	Assigned to	Due date
STEP IN (Designing and doing the evaluation)	9. Have you agreed on your evaluation questions? How are they linked to the components of the logic diagram representing the aims and services offered by your digital health community? How are they linked to measurement tools like surveys or analysis of CRM data?	8-14			
	10. What evaluation questions are of interest to whom? Are these realistic, given the project design and likely data availability? Can they be prioritised? How do people want to see the results used? Is this realistic?	8-9 and 21			
	11. Are there agreed evaluation indicators (criteria and standards) for each expected outcome?	6-10			
	12. Where evaluation outcome data is not yet available, do existing staff and systems have the capacity to do so in the future? Are responsibilities, sources and periodicities defined and appropriate? Is the budget adequate?	5-7			
	13. Are there data security risks? Describe them and propose a data management plan. Are the ethical issues known or knowable? Name them. Are they likely to be manageable? What constraints will they impose?	12-19			
	14. What is your analysis strategy? E.g., if you are collecting quantitative data, will you use descriptive statistics? How? Have you got advanced statistical knowledge in your team? If you are collecting qualitative data, will you use thematic analysis? How will you mix the results from quantitative and qualitative data analysis?	20			
STEP OUT (Using the evaluation findings)	15. Is there an opportunity for the evaluation to have an influence? Has the project accumulated enough experience/evidence to enable useful lessons? If the evaluation was planned in advance, is the evaluation still relevant?	5 and 7			

Appendix

1. Audrain-Pontevia, A.F. and Menvielle, L., 2018. Do online health communities enhance patient–physician relationship? An assessment of the impact of social support and patient empowerment. Health services management research, 31(3), pp.154-162.

2. BetterEvaluation (2021) What is Evaluation. https://www.betterevaluation.org/en/what-evaluation

3. This table is adapted from Mapping evaluation questions and indicators. Figure 3, p.23. Ontario Agency for Health Protection and Promotion (Public Health Ontario), Snelling S, Meserve A. Evaluating health promotion programs: introductory workbook. Toronto, ON: Queen's Printer for Ontario; 2016. ISBN: 978-1-4606-8317-0 https://www.publichealthontario.ca/-/media/documents/E/2016/evaluating-hp-programs-workbook.pdf?la=en

4. E. Smailhodzic, W. Hooijsma, A. Boonstra, D.J. Langley. Social media use in healthcare: A systematic review of effects on patients and on their relationship with healthcare professionals. BMC Health Services Research, 16 (1) (2016)

5. euroqol.org

6. sheffield.ac.uk/scharr/research/themes/valuing-health#SF-6D

7. Bowling, A., 2005. Mode of questionnaire administration can have serious effects on data quality. Journal of public health, 27(3), pp.281-291. https://academic.oup.com/jpubhealth/article/27/3/281/1511097?login=true

8. Adapted and simplified from Davies R (2013). Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations. DFID Working Paper 40. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/248656/wp40-planning-eval-assessments.pdf

UNIVERSITY OF LEEDS

Shift.ms