Epidemiology: Open Acess

Review Article | Open Access

SMS Phone Surveys and Mass-Messaging: Promises and Pitfalls

Havard Wahl Kongsgard[1,3*] , Tore Syversen[2] and Steinar Krokstad[1,4]

[1]HUNT Research Centre, Department of Public Health and General Practice, Faculty of Medicine, Norwegian University of Science and Technology, Norway

[2]Department of Neuroscience, Faculty of Medicine, Norwegian University of Science and Technology, Norway

[3]Department of Mathematical Sciences, Faculty of Information Technology, Mathematics and Electrical Engineering, Norwegian University of Science and Technology, Norway

[4]Levanger Hospital, Nord-Trondelag Health Trust, Norway

[*]**Corresponding author:** Havard Wahl Kongsgard, HUNT Research Centre, Department of Public Health and General Practice, Faculty of Medicine, Norwegian University of Science and Technology , Norway, Tel: 0047-91767991; E-mail: haavard.kongsgaard@gmail.com

Abstract

Data collection is often tedious and costly. Collecting data with mobile phones and SMS messages is one of many alternatives to more traditional methods. Mass messaging using SMS is a fast and cheap way to initiate a new study. However, as these data collection methods differ widely from conventional surveys, there are many pitfalls. This paper discusses best practice for SMS surveys and mass-messaging as a survey method. We argue that the best approach is to keep everything simple, invite a great number of respondents and know when to give up. Using SMS's or SMS systems for population wide data collection is a form of "mass messaging". Some European countries enforce an all-out ban on all SPAM like communication or require prior-consent. However, non-commercial actors like researchers are often accepted from such regulations. While plain text messages might be regarded as old and unattractive, the technology is fairly robust as it is very simple to generate and use. All types of mass-messaging give a low response-rate as e.g. email SPAM will generate typical response rates in the extent 1 of 12,500,000. But since scaling the survey will generate thousands of observations, low response-rate is not the main problem. Instead given the nature of mass-messaging, selection bias and data quality is a major concern. Nevertheless, these issues are not new to medicine and should not come as a surprise to anyone.

Keywords: SMS; Mobile phone; Mass-messaging; Method; Survey

Introduction

Data collection is often tedious and costly. Collecting data with mobile phones and SMS messages is one of many alternatives to more traditional methods. Mass-messaging using SMS is a fast and cheap way to initiate a new study. However, as these data collection methods differ widely from conventional surveys, there are many pitfalls. This paper discusses best practice for SMS surveys and mass-messaging as a survey method. We argue that the best approach is to keep everything simple, invite a great number of respondents and know when to give up.

Using SMS's or SMS systems for population wide data collection is a form of "mass-messaging". Some European countries enforce an all-out ban on all SPAM like communication or require prior-consent [1]. However, non-commercial actors like researchers are often excepted from such regulations [2]. While plain text messages might be regarded as old and unattractive, the technology is fairly robust as it is very simple to generate and use.

All types of mass-messaging give a low response-rate as e.g. email SPAM will generate typical response rates in the extent 1 of 12,500,000 [3]. But since scaling the survey will generate thousands of observations, low response-rate is not the main problem. Instead given the nature of mass-messaging, selection bias and data quality is a major concern. Nevertheless, these issues are not new to medicine and should not come as a surprise to anyone.

Creating a Survey

For SMS messages the simplest approach is to send a link to an online survey. Given the wide-spread use of internet enabled smartphones, websites are now more accessible than ever. However, most users will never complete the survey. In one experiment we found that while 3,000 of 20,000 accessed the website directly with their phone, only 500 completed the survey. Most phones will automatically generate a HTTP link, so the problem is not that people do not bother to type the web address. Instead we believe that since conventional surveys are not perceived as very interesting, people simply do not take the time to complete them.

An alternative is a "SMS system", a two-way text service where the participants communicate via text messages. The reason why it often works better, users are engaged in a different way. At work is essentially the same mechanism as in "gamification" of services, the use of game thinking and game mechanics in a non-game context to engage users and solve problems [4]. The bottom line, the way users interact with a service can substantially impact its popularity. In particular if users find something new and interesting, they are more likely to take part. Another benefit, SMS are less intrusive than automated phone survives, which in turn can lead to a better response rate

Smartphone apps are another option, but end-users dislike complicated procedures like downloading and installing apps, so the response-rate is likely to be very low. Take for example "The Human Face of Big Data", a large industry sponsored project. Their aim was to collect data from millions of users with an app [5]. In the end it got

little response despite wide media coverage and a large budget. Thus, smartphone apps seems to be a bad idea for collecting population wide data.

The Technical Side

A straightforward approach is to use a SMS gateway together with a database and a HTTP server. In most cases the SMS gateway is installed on a regular smartphone (Figure 1). Online SMS survey systems are available, but creating a system is strait forward and gives the researcher more control [6]. However, before create your own system, keep in mind that failure will be an option, so make sure that as much as possible can be reused in other projects.

Figure 1: Example of a Simple SMS System
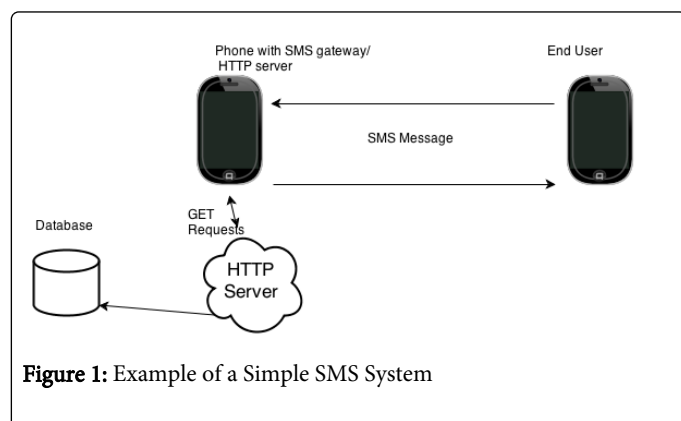
Survey

Feed the questions in sequential order or based on prior response. In order to avoid resending questions use an elegant question topology, for example a tree or directed acyclic graph (DAG) (Figure 2).

Figure 2: Hierarchical question layout, with DAG & tree structure
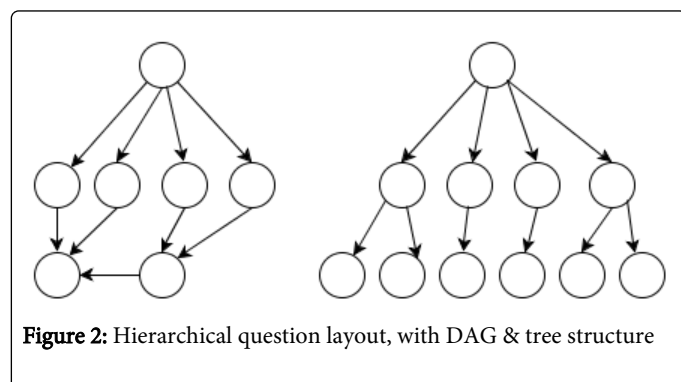
It's more time consuming to answer questions via text messages than on a website, so don't ask too many questions, a maximum of 2 to 5 questions is enough. It's often best to ask only simple and general question, but at the same time make it interesting and fun. As much information is available via the phone number, it's best to assume that this data is correct than to ask the person to verify such information.

Questions regarding current debate or news stories in general will get a high response. However, these questions leaves you with a narrow time frame, so in order to succeed you need a rapid deployment of the survey.

Text based responses can be difficult to handle, so keep all valid responses numeric. Avoid setting constraints on responses, instead terminate the survey with a standard response, or feed the user the next question. Test everything thoroughly in advanced, and try to find every "bug" before the participants. Remember that if something goes wrong the first time, the average user will not return for a second attempt.

Some participants equal medical surveys to small black holes. Too often participants will never get to know the result, what the data was used for or if the entire thing was a waste of time. To personalize the study and encourage participation in the future, consider sending a feedback to participants, summarizing the results of the study.

While the care free nature of such studies is a major advantage, it's also a major problem. In theory it's possible to conduct all types of studies, but the nature imply favoring studies that do not demand much of the participants. For example, it would be tempting to use mass-messaging to promote conventional studies, involving blood venipuncture and other clinical measures. However, we suspect that the majority of responders would not turn up for the main part of the study. And overall we fair that a lack of personal commitment might lead to a lower turn out in total, compared to other approaches with a fewer responses in total.

Sample, Segment and Bias

Unlike regular phones, mobile phone numbers are personal. With public data linked to the number it is easy to preselect population segments, for example only people in polluted areas, a specific age group or former patients at a hospital. However, with mass-messaging pre-sampling is problematic. The response rate is low, often less than 5 percent. In some cases even lower as the responders can be outright hostile confusing the survey with SPAM. The main priority should always be to get as many responses as possible. As some groups obviously can have a high response on specific topics, take care when pre-sampling participants. Worse, in some settings special subgroups make up the majority of responders, while they only represent a fraction of the population.

Scaling the survey and sample from the collected data is the best approach for dealing with these issues. For general inquiries we recommend a minimum of 100 000 invitations. While 5 percent response-rate might seem low, take into consideration the absolute number of participants. Most health survives have less than 6 000 participants, in comparison with a good design it's possible to get over 50 000. In many settings, increasing the sample size is enough to ensure an independent and representable sample. In practice however, the population is limited and the collected data often biased. For data collected with SMS's you should assume the latter is the case.

The 5 Hour Rule

As with most launches, the first hours are critical. For example on the crowdfunding website kickstarter.com, the initial response after 4 hours is enough to predict if a project fails or succeeds [7]. For SMS messages, if a user has not responded within 5 hours, they most likely never will. Our experience is that resending invitations more than once will not improve the overall response rate substantially.

In our case the optimal setup was to send out invitations in small batches, 200 – 1500 numbers with a 4 - 6 hours interval. Batches give researchers new possibilities. Unlike conventional survives, it's possible to alter the survey along the way. Big IT firms like Facebook &

Google often test new features on random users before altering their main content [8]. With SMS's you can do something similar, experiment with different samples/questions, and then quickly measure the response on different batches.

The Right Attitude

Lastly, having the right attitude matters more than anything else. Although SMS surveys are easy to implement, poorly planned surveys perceived negatively in the population and can undermine all kinds of health surveys. The response-rate is unpredictable and the data often biased, so there is no way of telling if the project will be a success or not. Therefore you have to learn how to be resilient and accept defeat before it's too late. After all, once you and your research team have designed a functional SMS system, the system can be reused in a new project.

Conclusion

Mass-messaging will never replace conventional surveys. These methods are cheap, fast and simple alternatives to conventional methods, when the conventional approach is unfeasible, too costly or simply too slow. However, using SMS systems as a research tool is not a walk in the park. Even for experienced researchers there are many pitfalls. Still the most important rules; keep it simple, scale the survey, experiment and know when to give up.

References

1. http://mailchimp.com/resources/guides/sms-spam-laws-and-lawsuits/html/.
2. http://lovdata.no/dokument/SF/forskrift/2004-02-16-401.
3. http://www.techradar.com/news/computing/spammers-get-1-response-to-12-500-000-emails-483381?src=rss&attr=all.
4. Easley, David, Jon Kleinberg (2010) Networks, Crowds, and Markets: Reasoning About a Highly Connected World, Cambridge University Press.
5. http://humanfaceofbigdata.com/.
6. http://textit.in.
7. http://vincent.etter.io/publicatins/etter 2013 cosn.pdf.
8. http://mashable.com/2013/08/07/facebook-promoted-trends/.