RESEARCH ARTICLE

Generation Scotland participant survey on data collection
[version 1; peer review: 1 approved, 1 approved with reservations]

Rachel Edwards, Archie Campbell, David Porteous

1 Centre for Genomic and Experimental Medicine, University of Edinburgh, Edinburgh, City of Edinburgh, EH4 2XU, UK
2 MRC Human Genetics Unit, University of Edinburgh, Edinburgh, City of Edinburgh, EH4 2XU, UK
3 Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, City of Edinburgh, EH4 2XU, UK

Abstract

Background: Generation Scotland (GS) is a population and family-based study of genetic and environmental health determinants. Recruitment to the Scottish Family Health Study component of GS took place between 2006-2011. Participants were aged 18 or over and consented to genetic studies, linkage to health records and recontact. Several recontact exercises have been successfully conducted aimed at a) recruitment to embedded or partner studies and b) the collection of additional data. As the cohort matures in age, we were interested in surveying attitudes to potential new approaches to data collection and recruitment.

Methods: A ten-question online survey was sent to those participants who provided an email address.

Results: We report a high level of positive responses to encouraging relatives to participate, to remote data and sample collection and for research access to stored newborn dried blood spots.

Conclusions: The majority of current and prospective GS participants are likely to respond positively to future requests for remote data and sample collection.

Keywords
Generation Scotland, participant survey, attitudes to data collection

This article is included in the Generation Scotland gateway.
Corresponding author: David Porteous (david.porteous@igmm.ed.ac.uk)

Author roles:
- **Edwards R**: Data Curation, Formal Analysis, Investigation, Project Administration, Resources, Writing – Review & Editing;
- **Campbell A**: Data Curation, Software, Writing – Review & Editing;
- **Porteous D**: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006]. Collection of email addresses was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award "STRatifying Resilience and Depression Longitudinally" (STRADL) Reference 104036). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2019 Edwards R et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Edwards R, Campbell A and Porteous D. Generation Scotland participant survey on data collection [version 1; peer review: 1 approved, 1 approved with reservations] Wellcome Open Research 2019, 4:111 (https://doi.org/10.12688/wellcomeopenres.15354.1)

Introduction
The Generation Scotland: Scottish Family Health Study (hereafter GS) cohort comprises nearly 24,000 participants in around 7,000 family groups, aged 18 or over at the time of recruitment. GP research practices in Glasgow, Tayside and Aberdeen provided letters of invitation on behalf of the study team. Participants were asked to recruit at least one family member to the study. Each recruit then completed a questionnaire on medical history, personality and lifestyle, attended a clinical examination, provided biological samples (blood and urine for genetic and biomarker studies) and consented to linkage to the routine medical records collected by NHS Scotland. We also asked permission to recontact participants for follow-on studies and 98% agreed to this. Participants are able to withdraw at any time. A small number of participants have withdrawn consent to recontact, but none have fully withdrawn to date.

Recruitment in 2006–2011 was limited by the funding available, not the willingness to participate. From information given at the time of recruitment and through maternity records, we know of a large number of first and second degree relatives of current participants that would add value to the cohort if interested and able to participate (estimated as >14,000). There are also significant numbers of younger relatives who might likewise be interested to participate.

Our research to date has shown the value of the family design to address questions that are beyond the easy reach of much larger cross-sectional studies, such as UK Biobank. These include spousal and household effects as well as parent-of-origin (genetic imprinting) effects and, of course, transgenerational effects.

When originally planning our recruitment to GS, we were conscious of the added value of birth data and sought to capture this where possible by recruiting in the Tayside region to co-capture members of the Walker Cohort and, at the tail end of recruitment, the Aberdeen Children of the Nineteen Fifties (ACONF).

Since the end of recruitment, internet access and smart phone use has increased dramatically across all ages and demographics. We have used the recontact mechanism to seek permission to contact participants by email – a third responded and agreed.

We conducted various public engagement events prior to recruitment to GS. We have also collected informal and formal feedback on participant attitudes in the past through various engagement events. Here, we describe a pilot study of attitudes to new participant recruitment and new modes of data collection. For simplicity, we used the popular SurveyMonkey tool to conduct the study, posing ten questions to guide our future planning.

Methods
Participants were recruited to GS via GP practices in Scotland and originally by letter. Email addresses were not collected at the time of recruitment but were asked for in a subsequent recontact exercise (STRADL), with 46% responding and 36% providing an email address. All those who shared an email address were eligible to take part in this questionnaire study. No other efforts to address potential bias were taken.

Our questionnaire, created using SurveyMonkey (San Mateo, California, USA), was entitled ‘Share your thoughts on health research’, was sent out between 3rd – 10th April 2019. The introduction to the questionnaire is shown in Box 1. The survey was sent to all 7,118 GS participants for whom we had email addresses, of which 2,613 responded (34%), the majority (76%) within a week on emailing. The time taken to complete the questionnaire averaged 2 minutes. There was no repeat emailing to non-respondents.

We started by asking about their age, gender and cohort participation. Cohort participation was queried so that the same questionnaire could be used for other cohorts and/or the general public. Here, we report only on the GS participants. We asked seven questions to explore current participant views on new participant recruitment and modes of data collection (Table 1). Questions 4–10 were formatted in the same five-option style: Definitely yes; Probably yes; Not sure; Probably no; Definitely no. We used a panel of experts in public engagement and Patient and Public Involvement panel members to check the wordage and formatting.

The demographics are shown in Figure 1–Figure 3 and results of the survey questions in Figure 4–Figure 10.
Table 1. ‘Share your thoughts on health research’: survey questions.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Question Asked</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What age are you?</td>
</tr>
<tr>
<td>2</td>
<td>What gender do you identify as?</td>
</tr>
<tr>
<td>3</td>
<td>Have you taken part in any of these health research studies? (Tick all that apply)</td>
</tr>
<tr>
<td>4</td>
<td>If asked by researchers at a University to take part in a health study like Generation Scotland, how would you answer?</td>
</tr>
<tr>
<td>5</td>
<td>If the study asked you to invite another member of your family to join too, would you be happy to do so?</td>
</tr>
<tr>
<td>6</td>
<td>If the study asked you to give details about your own health, personality, habits, and lifestyle, would you be happy to do so?</td>
</tr>
<tr>
<td>7</td>
<td>If the study asked to use blood samples left over from routine health tests, would you agree to that?</td>
</tr>
<tr>
<td>8</td>
<td>If the study involved genetic analysis, using DNA extracted from a saliva sample kit sent to your home, would you agree to take part?</td>
</tr>
<tr>
<td>9</td>
<td>If the study involved analysis, using blood extracted from a finger prick kit sent to your home, would you agree to take part?</td>
</tr>
<tr>
<td>10</td>
<td>Children born in Scotland after 1965 have heel prick blood spots (Guthrie cards) taken to test for serious but treatable conditions. The Guthrie cards have been kept in a safe place by the NHS. A study might ask for specific permission to use these stored blood spots for research. Does this sound reasonable, whether or not your own Guthrie card has been stored?</td>
</tr>
</tbody>
</table>

Figure 1. What age are you?

Ethical issues
Generation Scotland is an approved research tissue bank. It was approved by East of Scotland Research Ethics Service (EoSRES) under reference LR/15/ES/0040. GS participants consented to recontact when they joined the Generation Scotland study. Participants who previously provided an email address were invited to take part via email, with participation in the survey being entirely voluntary. GS participants are free to withdraw at any time.

Results and discussion
There were more female (1,660, 64%) than male (944, 36%) respondents, but this reflects in part the baseline participation (59% female:41% male) and those sharing email addresses (61% female, 39% male). Respondents were consistently positive in answers to questions 4–10, with over 80% replying ‘definitely yes’ or ‘probably yes’. The definite ‘no’ category was returned by 1% or less of respondents. Individual and summary level results of the survey are available as Underlying data15.
The questions that sought opinions on the acceptability of potential future means of recruitment, data and sample collection are shown in Table 1. Question 4 was posed because the original recruitment was by a letter of introduction by their GP. It was encouraging that recruitment via researchers at a University met with a strongly positive response. Question 5 was posed because our GS research to date has highlighted the value of family structure, and both genetic and non-genetic (spousal, household, environmental) effects on health and health trajectories\textsuperscript{1-6}. Question 6 was posed because this type of information was included in the original pre-clinic questionnaire and has been of great value in addressing questions relating to mental health, cognition, personality and lifestyle on physical health, and vice versa\textsuperscript{1,4,6,16,17}. Question 7 was posed because this could
be a very cost-effective way of collecting blood samples (and potentially other left-over pathology samples). In Scotland, the SHARE mechanism is already in place to consent to left-over blood being made available for research\textsuperscript{11}. Over 250,000 have already signed up at https://www.registerforshare.org/.

Questions 8 and 9 were posed because clinic-based (or GP practice based) recruitment is costly and can exclude or limit participation for geographical reasons. By contrast, online shopping and postal delivery of a wide range of products, including medical products and prescriptions, is now commonplace. Home-based
kits for saliva and blood samples are well tried and tested. Blood prick sampling has been the norm for blood glucose measurements in diabetes for many years and is well tolerated.

Question 10 sought opinions on research access to archived new-born dried blood spots. Heel prick blood sampling in new-borns is standard practice in many countries and used to screen for a well-defined set of rare and preventable metabolic disorders and/or detection of genetically inherited conditions that benefit from early detection. These dried blood spots can be used for a variety of other tests, including infection, toxins and epigenetic modification (DNA methylation). In many countries,
Figure 8. Would you agree to join if asked to provide a saliva sample for genetic analysis?

Figure 9. Would you take part if asked to complete a home finger prick blood test?

these cards are held long-term. In Scotland, new-born heel prick dried blood samples (historically referred to as Guthrie cards) have been collected and retained since 1965, 3 million to date, increasing by around 60,000 per annum. We estimate that up to 9,700 (40%) of current GS participants will have a blood spot card retained and stored securely by the NHS Scotland National Screening Service. In some countries, such as Denmark, research access is possible under agreed conditions and for agreed purposes\textsuperscript{15}. Whilst this is not currently approved in the UK, we felt it important to test current opinion amongst medical
Would you agree to your Guthrie card being used for analysis? Although 1% of respondents said definitely ‘no’ to research use of their stored dried blood spot card on principle, over 90% said definitely or probably ‘yes’.

In summary, we found that there was a very positive overall response to questions designed to assess the likely ‘buy in’ to a cost-effective extension to the Generation Scotland cohort. The high response rate suggested that this easy-to-offer approach could be a useful tool for further engagement and shaping of cohort retention, expansion and enhancement. One caveat to the wider interpretation of our results is that we are consulting existing participants, not the general public or hard-to-reach constituents. Further comparative contact studies are warranted. For the practical purpose however of scoping the possible expansion and extension of GS, these findings are valuable and encouraging.

**Data availability**

The SurveyMonkey study was by pseudo-anonymised email recontact. All summary and individual level survey responses to this can be found in DataShare.

**Underlying data**

Edinburgh DataShare: Generation Scotland Survey Monkey data. [https://doi.org/10.7488/ds/2585](https://doi.org/10.7488/ds/2585). This project contains the following underlying data:

Generation Scotland Survey Monkey data.xlsx (includes all summary and individual level data collected in the survey).

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Researchers can request access to this and all other GS data by contacting the GS Access Committee. A phenotype data dictionary is available and open access GWAS summary statistics can be downloaded. Non-identifiable information from the GS-SFHS cohort is available to researchers in the UK and to international collaborators through application to the GS Access Committee. GS operates a managed data access process including an online application form, and proposals are reviewed by the GS Access Committee. Summary information to help researchers assess the feasibility and statistical power of a proposed project is available on request by contacting resources@generationscotland.org.

**Grant information**

Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006]. Collection of email addresses was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award “STRatifying Resilience and Depression Longitudinally” (STRADL) Reference 104036).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
Acknowledgments

We are grateful to all the families who took part in Generation Scotland, the general practitioners and the Scottish School of Primary Care for their help in recruiting them, and the whole Generation Scotland team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, healthcare assistants and nurses.

References

Open Peer Review

Current Peer Review Status: ✔️ ❓

Version 1

Reviewer Report 26 November 2019

https://doi.org/10.21956/wellcomeopenres.16769.r36432

© 2019 Steves C. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Claire J. Steves
The Department of Twin Research and Genetic Epidemiology, St Thomas’ Hospital, Kings College London, London, UK

This is a clearly written report surveying 36% of Generation Scotland participants who have given their email addresses on their views on extension of the study. The response rate was 34%. Generally the responses are very positive about planned initiatives. I would be keen for the report to also include information (e.g. a table) on the demographic characteristics of the non-responders, and those who did not have or did not share their email, to assess what bias either of these sub-setting processes might have had. It is likely that responders may be slightly more likely to engage with the initiatives planned. Provided the survey responses can be linked to the demographics of the participants collected at enrolment, more detail on the characteristics of the ~1% less willing subjects, together with the numbers with these characteristics in the whole cohort could be informative of the bias which might occur if consent for these initiatives is sought.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
I am Deputy Director for TwinsUK which is another UK family cohort. We are working to share good practice with Generation Scotland on participant engagement and collection methods. I do not think this has affected my ability to provide an objective review.

**Reviewer Expertise:** Longitudinal cohort studies.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 07 October 2019

https://doi.org/10.21956/wellcomeopenres.16769.r36433

© 2019 Douglas E. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Elaine Douglas
University of Stirling, Stirling, UK

Scotland has a rich data infrastructure which makes it well placed to conduct biomedical research that can be linked to individual survey results and administrative data. The potential of such infrastructure can only be achieved by the trust and consent of the public to take part in such research. This paper provides a welcome overview of the willingness of participants to take part in such research. I recommend this paper for indexing.

Comments

- While the sample are drawn from existing GS participants and therefore not generalisable to the wider population per se, the response mode and response rates are relevant, and will be of interest to, other researchers who may be considering a follow-up to their own participants.

- The results of this research e.g. the willingness of the sample to take part in health studies, to support recruitment via family members, and to give consent to use of blood samples, etc. is of value to researchers working within this field, and gives some insight (with the aforementioned caveats) of the acceptability of such research and recruitment methods.

Suggestions:

- If the y-axis of all figures were produced using the same scale this would ease the comparison of responses across all questions.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes
Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Survey design and implementation, administrative data linkage, ageing, health behaviour, public health

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

---

Author Response 22 Oct 2019

**David Porteous**, University of Edinburgh, Edinburgh, UK

Thank you for these comments. We will provide a revised submission with all of the questionnaire responses on the same scale. We have not changed any of the text.

**Competing Interests:** None