Recontacting biobank participants to collect lifestyle, behavioural and cognitive information via online questionnaires: lessons from a pilot study within FinnGen

Rodosthenous S Rodosthenous,1, Mari E K Niemi,1 Lila Kallio,2 Merja Perala,2 Perttu Terho,2 Theresa Knopp,3 Eero Punkka,3 Enni M Makkonen,4 Paula Nurmi,5 Johanna Makela,5 Pauli Wihuri,6 Marco Hautalahti,6 Corianna Moffatt,7 Paolo Martini,7 Laura Germine,7 Viola A Makela,1 Oona A Karhunen,1 Jari Lahti,8 Tero S Hiekkalinna,5 Tero Jyrhama,1 Huei-yi Shen,1 Heiko Runz,9 Aarno Palotie,1,10 Markus Perola,5 Andrea Ganna,1 FinnGen

ABSTRACT

Objectives To recontact biobank participants and collect cognitive, behavioural and lifestyle information via a secure online platform.

Design Biobank-based recontacting pilot study.

Setting Three Finnish biobanks (Helsinki, Auria, Tampere) recruiting participants from February 2021 to July 2021.

Participants All eligible invitees were enrolled in FinnGen by their biobanks (Helsinki, Auria, Tampere), had available genetic data and were >18 years old. Individuals with severe neuropsychiatric disease or cognitive or physical disabilities were excluded. Lastly, 5995 participants were selected based on their polygenic score for cognitive abilities and invited to the study. Among invitees, 1115 had successfully participated and completed the study questionnaire(s).

Outcome measures The primary outcome was the participation rate among study invitees. Secondary outcomes included questionnaire completion rate, quality of data collected and comparison of participation rate boosting strategies.

Results The overall participation rate was 18.6% among all invitees and 23.1% among individuals aged 18–69. A second reminder letter yielded an additional 9.7% participation rate in those who did not respond to the first invitation. Recontacting participants via an online healthcare portal yielded lower participation than recontacting via physical letter. The completion rate of the questionnaire and cognitive tests was high (92% and 85%, respectively), and measurements were overall reliable among participants. For example, the correlation (r) between self-reported body mass index and that collected by the biobanks was 0.92.

Conclusion In summary, this pilot suggests that recontacting FinnGen participants with the goal to collect a wide range of cognitive, behavioural and lifestyle information without additional engagement results in a low participation rate, but with reliable data. We suggest that such information be collected at enrolment, if possible, rather than via post hoc recontacting.

INTRODUCTION

Biobank studies are being set up across the world.1-9 These studies are characterised by the possibility to link biological measurements, such as DNA, proteins and metabolites, with extensive longitudinal health information. Health outcomes are often obtained by linkage with electronic health records or national registers, as in the case of the Nordic countries. Biobank studies have allowed extensive characterisation of the genetic architecture of common diseases,10-12 provided novel epidemiological insights13 14 and identified novel disease markers.15

The UK Biobank study,1,16 one of the largest and most widely used biobanks, has also collected lifestyle, behavioural information,
and anthropometric measurements for all their participants. This was possible, at the cost of collecting a non-representative sample of the population, because all participants were enrolled via in-person visit to one of the 22 recruitment centres. However, most of the other biobank studies use a different approach collecting samples via the healthcare system or via other approaches that do not entail an extensive in-person examination at recruitment. For this reason, it is often difficult to obtain extensive behavioural and lifestyle information from biobank participants and recontacting after recruitment is required.

The FinnGen study is a public–private partnership research project combining genotype data generated from Finnish biobank samples and digital health record data from Finnish health registers (https://www.finngen.fi/en) aiming to provide new insight in disease genetics. Up to 500,000 participants of Finnish ancestry will be part of FinnGen and >350,000 have already been genotyped and linked with comprehensive health registers. Participants to the FinnGen study are recruited by several biobanks across Finland and all participants have signed a broad biobank consent in accordance with the Finnish Biobank Act. Participants are enrolled because they are part of previous research studies or via hospitals and blood donation centres, but no extensive behavioural and lifestyle information is systematically collected for everyone at recruitment. Overall, FinnGen does not suffer from the ‘healthy volunteer effect’ and, on the contrary, is enriched for individuals who are more likely to have been in contact with the healthcare system.

Study population

We used data from FinnGen release R6, which included 259,578 individuals of Finnish ancestry and genetic information available. According to the Finnish Biobank Act (688/2012), biobanks may recontact a person who has given such permission in his/her biobank consent. Three biobanks (Helsinki, Auria and Tampere) encompassing 100,040 FinnGen participants (Helsinki n=58,518; Auria n=29,159; Tampere n=12,363) participated in the pilot study. We included individuals who were 18 years or older at the time of the initiation of the study (February 2021). For Helsinki Biobank, we restricted inclusion to individuals younger than 70 years of age. We excluded individuals with severe neuropsychiatric disease, or cognitive or physical disabilities, as we expected the participation rate to be lower in this group. More specifically, we excluded individuals with any of the following diagnoses as obtained from the health register data: (1) any dementia (International Classification of Diseases 10th Revision (ICD-10): G30, F051, F00–03); (2) neurodegenerative diseases (ICD-10: G310–312, G318–319); (3) Parkinson’s disease (ICD-10: G20); (4) multiple sclerosis (ICD-10: G35); (5) schizophrenia, schizotypal and delusional disorders (ICD-10: F20); (6) mental retardation (ICD-10: F70–73, F78–79); (7) stroke (ICD-10: I60–64); (8) transient ischaemic attack (ICD-10: G45); (9) visual impairment including blindness (binocular or monocular) (ICD-10: H54). Finally, we selected individuals in the top and bottom 3% of a polygenic score (PS) for cognitive performances. This was motivated by previous work showing a strong association between a PS for cognitive performances and study participation. Our aim was to prospectively assess if recruiting individuals based on this PS would result in a different study participation.

METHODS

Study population

We used data from FinnGen release R6, which included 259,578 individuals of Finnish ancestry and genetic information available. According to the Finnish Biobank Act (688/2012), biobanks may recontact a person who has given such permission in his/her biobank consent. Three biobanks (Helsinki, Auria and Tampere) encompassing 100,040 FinnGen participants (Helsinki n=58,518; Auria n=29,159; Tampere n=12,363) participated in the pilot study. We included individuals who were 18 years or older at the time of the initiation of the study (February 2021). For Helsinki Biobank, we restricted inclusion to individuals younger than 70 years of age. We excluded individuals with severe neuropsychiatric disease, or cognitive or physical disabilities, as we expected the participation rate to be lower in this group. More specifically, we excluded individuals with any of the following diagnoses as obtained from the health register data: (1) any dementia (International Classification of Diseases 10th Revision (ICD-10): G30, F051, F00–03); (2) neurodegenerative diseases (ICD-10: G310–312, G318–319); (3) Parkinson’s disease (ICD-10: G20); (4) multiple sclerosis (ICD-10: G35); (5) schizophrenia, schizotypal and delusional disorders (ICD-10: F20); (6) mental retardation (ICD-10: F70–73, F78–79); (7) stroke (ICD-10: I60–64); (8) transient ischaemic attack (ICD-10: G45); (9) visual impairment including blindness (binocular or monocular) (ICD-10: H54). Finally, we selected individuals in the top and bottom 3% of a polygenic score (PS) for cognitive performances. This was motivated by previous work showing a strong association between a PS for cognitive performances and study participation. Our aim was to prospectively assess if recruiting individuals based on this PS would result in a different study participation.

Invitation procedures

Invitation was carried out via postal letter for 5995 FinnGen participants by Helsinki Biobank (n=2000), Auria Biobank (n=2000) and Tampere Biobank (n=1236). For an additional 759 individuals from Tampere Biobank, an invitation was sent electronically via the OmaTays healthcare portal, which has been used already for medical communication and online booking appointments by the Tampere Biobank. A random subset of 422 participants aged 18–69 from Tampere Biobank (n=240) and Auria Biobank (n=182) who did not reply to the first invitation were recontacted with a second invitation letter.

Portal for enrolment and collecting survey information

We collaborated with the Finnish Biobank Cooperative to create a portal that allows study participants to securely log in, identify themselves, provide consent and access the study survey tools. The portal is hosted at https://omabiopankki.fingenious.fi/ and secure authentication is guaranteed by a bank ID identification system which is available to most Finnish residents. Once the participants log into the system, the study invitation and consent can be viewed online. After the participants consent to participate in the study, they can view the two information collection tools: a general questionnaire and a battery of the cognitive tests. By selecting each tool, participants first view a short description prior to their completion. Both the general questionnaire and the cognitive tests are developed on third-party platforms. Once the participants have completed the online questionnaire or cognitive test battery, they are redirected back to the OmaBiopankki portal where they can complete the remaining tasks.

General study questionnaire

The general study questionnaire included 18 categories of questions which cover a broad range of health-related topics that are not obtainable or are difficult to obtain through hospital and register data. All questions were accompanied by the options ‘don’t know’ and ‘prefer not to answer’ to allow the participants to voluntarily skip any
of the questions. The following categories were included, with the scientific justification: (1) general questions (sex at birth, height, weight); (2) women’s health; (3) smoking; (4) family medical history; (5) disease diagnoses (including diseases not well captured by hospital and register data); (6) medication history (including over-the-counter medications); (7) alcohol consumption; (8) early-onset neurodevelopmental disorders and psychiatric disorders; (9) mental health and mood; (10) education levels; (11) physical activity; (12) multisite pain; (13) risk taking; (14) influenza and viral infections (respiratory); (15) sleep; (16) oral health; (17) diet; (18) sauna habits.

Cognitive tests
We collected data from a test battery designed to capture different aspects of cognitive abilities. This test battery was provided by TestMyBrain and translated into Finnish. TestMyBrain has previously validated all their cognitive test batteries (https://psyarxiv.com/dcszr/). Eight different tests were performed: digit symbol matching, flicker change detection, visual paired associates, multi-racial emotion identification, gradual-onset continuous performance, matrix reasoning, verbal and vocabulary test. Because the vocabulary test could not be directly translated from English, we created a Finnish-specific version of the test by selecting 19 out of 30 candidate words that provided with the highest correlation with the well-established Wechsler Adult Intelligence Scale-Revised (WAIS-R) vocabulary task in a sample of n=24 Finns (79.2% women) with an average age of 24.3 years (SD: 3.0). Our new online vocabulary test had a Spearman’s correlation of 0.86 with the vocabulary task from the WAIS-R. The analysis of collected cognitive test data is out of the scope of this paper and therefore not shown here. Participants could see a summary of their results and how they scored in each test compared with the average scores after completion of all the tests.

Feedback questionnaire
To get a sense of the participants’ experience on recontacting, we introduced a feedback questionnaire directed to the individuals logging into the OmaBiopankki portal. The main question to participants was to rate their overall experience in completing the online questionnaire and cognitive tests using a Net Promoter Score (NPS) scale from 0 to 10. In addition, we asked if they have participated in an online questionnaire study before, if they experienced any technical issues while completing the survey and to share with us what they enjoyed or did not enjoy while completing the survey.

Polygenic risk score for cognitive abilities
Summary statistics used to generate the PS were obtained by combining results from two highly correlated genome-wide association studies (GWAS): educational attainment and intelligence using a multtrait analysis of GWAS method as described elsewhere. The PS for cognitive abilities was then generated based on these summary statistics using a pruning and thresholding approach including all single nucleotide polymorphisms (SNPs) with a p value <1. The decision of selecting individuals based on genetic scores was taken to maximise the power of identifying association between genetic information and cognitive domains. Selecting individuals from the top and bottom 3% of the PS distribution was informed by the powder calculations, study design and budget constraints.

Statistical analysis
A Wilcoxon rank-sum test was performed to test (α=0.05 significance level) whether the median age and body mass index (BMI) were significantly different between invitees who participated or not. To test the difference in other categorical variables (eg, education and disease prevalence) between the two groups, we used the two-sample proportions test (α=0.05 significance level). All descriptive and statistical analyses were performed in R V.4.2.1.

Patient and public involvement
Participants were not involved in the design, choice of outcome measures or recruitment and conducting strategy in this study. All participants were given the option to contact the research team and ask any questions both prior and during the study. All study participants were sent a questionnaire to provide their feedback about the study.

RESULTS
In this genetically informed pilot study, we invited 5995 FinnGen participants (66% female) across three biobanks who were selected according to the inclusion criteria shown in figure 1. Out of 5995 individuals invited, 1115 (73% female) accessed the study online portal and successfully completed the general questionnaire with an overall participation rate of 18.6% (table 1). The participation rate in the 18–69 years old age group was slightly higher at 23.1% (1018/4399). The highest overall participation rate across all ages was observed in Helsinki Biobank with 23%, followed by Tampere Biobank with 17% and Auria Biobank with 15.8%. Lower participation rate in Tampere Biobank and Auria Biobank is partially explained by the higher age range of the invited individuals. Among the 18–69 years old range, the participation rate was similar across all biobanks. The highest participation rate was consistently observed in the 40–69 years old age group across all biobanks (table 1 and online supplemental table 1). The participation rates among individuals who were ≥70 years old were substantially lower (6.1%) as compared with the <70 years old group (23.1%). We also tested if sending a follow-up reminder letter would significantly improve participation rate among non-responders. To assess this, we sent a reminder by post to 422 individuals from Tampere Biobank (n=240) and Auria Biobank (n=182) who were
invited initially but did not participate in the study during the first phase. The overall response rate for this second letter was 9.7% (41/422) higher in Tampere Biobank (11.5%, 28/240) than Auria Biobank (7.1%, 13/182).

Using FinnGen data, we further investigated whether any differences in basic characteristics and disease prevalence among invitees could have potentially impacted their participation in the study. No significant age difference was observed between invitees who participated (54.9±13.5 (mean±SD)) versus those who did not participate (55.6±15.6 (mean±SD)) in the study (p=0.08); however, we found that invitees who participated had a lower BMI (28.0±6.2 (mean±SD)) compared with those who did not participate (28.5±6.2 (mean±SD)) in the study (p=0.02). We also found that invitees with a university degree or higher were more likely to participate (24%, 433/1810) as compared with invitees who had a lower education (17%, 519/3099). The difference in the two proportions between the groups was statistically significant (p<0.0001). When assessing health information using FinnGen data, we found that a significantly higher (p<0.0001) proportion of invitees who did not participate were previously diagnosed with hypertension (24.9%, 1210/4854) as compared with invitees who participated (16.9%, 186/1101). No differences in other disease prevalences, such as asthma, arthrosis, depression and immune bowel disease, were observed between the two groups.

In addition, we compared the participation rate between individuals invited via physical letters and those invited electronically via the Tampere Biobank OmaTays healthcare portal. Out of 1995 invitees, 759 received their invitations via the OmaTays healthcare portal and 1236 via physical letter. The participation rate was 12.1% among those invited via the OmaTays compared with 21.4% among those who were invited via a physical letter from Tampere Biobank. Data retrieved from the OmaTays healthcare portal showed that 451/759 (59.5%) of those who were invited electronically did not open the invitation at all. Among the rest who viewed the invitation, ~30% (92/308) went on to participate in the study. Individuals invited in this pilot study were consented in the first place via the biobank consents, in accordance with the Finnish Biobank Act. Thus, they could decide to withdraw their consent for their samples to be used in research studies. In this study, 0.3% of the invited individuals contacted any of the three biobanks to withdraw their consent.

Among individuals who started answering the online questionnaire, the completion rate was 92%. The completion rate for the entire battery of cognitive tests was 85%, despite the length of the tests (estimated to be 30–40 min).
However, substantially fewer people started the cognitive test compared with the general questionnaire (n=699 vs n=1115). Among 143 participants who provided their experience through the feedback questionnaire, the median NPS was 8, indicating that most individuals were happy with the current design. More specifically, 121/143 (85%) gave a score >7, which is interpreted as good/very good. On the contrary, 22/143 (15%) scored the questionnaire with a 6 or less, indicating they had a negative experience (figure 2). It is worth mentioning that only 53% (73/137) had previous experience with online questionnaires, whereas for 47% (64/137) of the participants this was their first online questionnaire.

For a subset of FinnGen participants, information about BMI was available and extracted from electronic health records or in-person visits. Thus, for 673 individuals, we could compare the self-reported BMI data obtained from the questionnaire with those previously available in FinnGen and extracted from electronic health records or in-person measurements. We found a high correlation between the two measurements (Spearman’s r=0.92) despite these being collected at different time points and with different approaches (figure 3). The mean±SD of the self-reported and FinnGen-collected BMI data was 28.1±6.0 and 28.0±6.2,
DISCUSSION

able and economical sample collection by using existing registries for cognitive abilities resulted in differential study participation. In particular, 19.4% of individuals who participated in this pilot study were in the top 3% of the PS for cognitive performance compared with 16.9% of those who did not participate. The difference in the two proportions between the two groups was statistically significant (p=0.01) but modest.

For some diseases, a higher prevalence is expected because health registers do not provide good coverage if diagnosed in primary care. On the other hand, some individuals might over-report or misreport disease diagnoses. A head-to-head comparison between self-reported and register-based diagnoses is challenging because different combinations of diagnostic and medication codes can be used to define the same disease from health registers.

Last, access to genetic data for both participants and non-participants in this pilot study allowed us to test if our strategy of recruiting individuals in the top and bottom 3% of a PS for cognitive abilities resulted in differential study participation between the two groups. We found that a higher PS for cognitive performance was associated with higher participation. In particular, 19.4% of individuals who participated in the pilot study were in the top 3% of the PS for cognitive performance compared with 16.9% of those who did not participate. The difference in the two proportions between the groups was statistically significant (p=0.01) but modest.

Several biobank studies are designed to prioritise scalable and economical sample collection by using existing biological banks or hospital-based recruitment strategies. Often, extensive characterisation of lifestyle, cognitive and behavioural information is done a posteriori by recontacting study participants. In this pilot study, we have proposed and tested a strategy to recontact participants in one of the largest biobank studies in the world: FinnGen. To this goal, we have established a scalable recontacting process and designed an online recontacting platform (ie, OmaBiopankki) for secure identification of participants. The platform can now be used as a benchmark for future recontacting studies in Finland.

Despite declining response rates in population-based surveys globally, research studies conducted in Finland have shown a high participation rate (>50%). In this pilot study, we have observed a lower participation rate of 23% in the 18–69 years old age range, which may be explained by several reasons. First, all participants provided a general consent to their biobank that covers current studies such as FinnGen and a wide range of possible future research studies. Therefore, consented individuals may not be aware of their participation or be directly engaged at the time of contacting on behalf of the FinnGen study. Second, this pilot study was designed to capture a broad range of cognitive, behavioural and lifestyle information from invited individuals and not to target any specific disease(s). The three biobanks included in this pilot recruit individuals who are hospitalised or have been directly in contact with the healthcare system. Thus, studies that target specific diseases directly relevant to each participant’s own health may result in higher participation rates. Third, for the same reason, consenting participants are likely to be sicker than the general population, which may subsequently impact their participation in such studies. In fact, lower participation rates among less healthy individuals are well documented.

As part of our efforts to boost the participation rate in this study, we tested several methods. Initially, we restricted the age group of invitees to under 70 years old because we observed a significantly lower participation rate among those over 70 years old. Based on the feedback we got through our study contact email and phone helpline, we believe that this difference was mainly due to the limited access or lack of familiarity with the use of the internet and mobile banking applications required for secure authentication among older individuals. In addition, we sent a follow-up letter as a reminder to a subset of non-responders to the initial invitation and observed a significant boost in response rate (9.7%). Our finding is very comparable to Harrison et al’s study that found a 9% increase in participation rates after sending a reminder letter to non-responders. Another study by Smith et al also found that sending follow-up letters significantly increases the chance of response among non-responders to the initial invitation.

We also evaluated response rates between sending a physical or an electronic letter and found that invitees who received a physical letter were much more likely to respond to the survey. In a recent study, researchers tested the same invitation methods and found a striking difference in participation rates when sending a physical (26.8%) or an online (1.8%) invitation.
To the best of our knowledge, this is one of the first studies that assessed participation rate in relation to the cognitive PS of participants. To further support our finding that individuals with a higher cognitive PS are more likely to participate in online surveys, we also compared the education level between those who participated compared with those who did not participate in the study. We found that invitees who had a university degree or higher were more likely to participate (24%, 433/1810) as compared with invitees who had a lower education (17%, 519/3099).

This study has also a few limitations. First, we could not assess the reasons for low participation because non-participants could not be further recontacted. Second, because this study could only recontact FinnGen participants, it essentially enrolled from a pool of individuals with a Finnish ancestry and a higher percentage of women than men. Therefore, generalising findings to other populations shall be done with caution.

In conclusion, this pilot suggests that recontacting individuals who have consented to be part of a biobank study with the goal to collect a wide range of cognitive, behavioural and lifestyle information can be challenging and may result in lower-than-expected participation. We speculate that returning some tangible incentive and/or relevant health information to participants might improve participation rates. Future studies are warranted to test this hypothesis and other strategies to improve participation in such survey research studies. Nonetheless, we suggest that cognitive, behavioural and lifestyle information be collected, whenever possible, at enrolment rather than via a post hoc recontacting process.

Author affiliations
1Institute for Molecular Medicine Finland (FIMM), HLIFE, University of Helsinki, Helsinki, Finland
2Turku University Hospital (TUKS), Turku, Finland
3Hospital District of Helsinki and Uusimaa, Helsinki, Finland
4Tampere University Hospital, Tampere, Finland
5Finnish Institute for Health and Welfare, Helsinki, Finland
6Finnish Biobank Cooperative, Helsinki, Finland
7The Many Brains Project, Boston, Massachusetts, USA
8Department of Psychology and Logopedics, Faculty of Medicine University of Helsinki, Helsinki, Finland
9Translational Biology, Research and Development, Biogen Inc, Cambridge, Massachusetts, USA
10The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

Acknowledgements We want to acknowledge the participants and investigators of FinnGen study. The following biobanks are acknowledged for delivering biobank samples to FinnGen: Auria Biobank (www.auria.fi/biopankki), THL Biobank (www.thl.fi/biobank), Helsinki Biobank (www.helsinginbiopankki.fi), Biobank Borealis of Northern Finland (https://www.psph.fi/Tutkimus-ja-opetus/Biopankki/Pages/Biobank-Borealis-briefly-in-English.aspx), Finnish Clinical Biobank Tampere (www.tays.ensu.fi/Research_development/Finnish_Clinical_Biobank_Tampere), Biobank of Eastern Finland (www.itu-suomenbiopankki.fi/en), Central Finland Biobank (www.kkshp.fi/fifi/PottiLaale Biopankki), Finnish Red Cross Blood Service Biobank (www.veripalvelu.fi/verenluovutus/biopankkitilointa) and Terveystalo Biobank (www.terveystalo.com/fi/Viisytietoa/Terveystalo-Biopankki/Biopankki/). All Finnish biobanks are members of BMBRI infrastructure (www.bbrmi.nl). The Finnish Biobank Cooperative (FBB) (https://fbb.fi) is the coordinator of BMBRI-ERIC operations in Finland. The Finnish biobank data can be accessed through the Fingenious services (https://site.fingenious.fi/en/) managed by FINBB.

Collaborators FinnGen: Adam Platt (Astra Zeneca, Cambridge, United Kingdom), Adnan Ziegelstein (University of Michigan, Ann Arbor, MI, United States), Adriana Guerrero-Elizalde (University of Miami, Miami, FL, United States), The Many Brains Project, Boston, Massachusetts, USA

This study has also a few limitations. First, we could not assess the reasons for low participation because non-participants could not be further recontacted. Second, because this study could only recontact FinnGen participants, it essentially enrolled from a pool of individuals with a Finnish ancestry and a higher percentage of women than men. Therefore, generalising findings to other populations shall be done with caution.

In conclusion, this pilot suggests that recontacting individuals who have consented to be part of a biobank study with the goal to collect a wide range of cognitive, behavioural and lifestyle information can be challenging and may result in lower-than-expected participation. We speculate that returning some tangible incentive and/or relevant health information to participants might improve participation rates. Future studies are warranted to test this hypothesis and other strategies to improve participation in such survey research studies. Nonetheless, we suggest that cognitive, behavioural and lifestyle information be collected, whenever possible, at enrolment rather than via a post hoc recontacting process.

Author affiliations
1Institute for Molecular Medicine Finland (FIMM), HLIFE, University of Helsinki, Helsinki, Finland
2Turku University Hospital (TUKS), Turku, Finland
3Hospital District of Helsinki and Uusimaa, Helsinki, Finland
4Tampere University Hospital, Tampere, Finland
5Finnish Institute for Health and Welfare, Helsinki, Finland
6Finnish Biobank Cooperative, Helsinki, Finland
7The Many Brains Project, Boston, Massachusetts, USA
8Department of Psychology and Logopedics, Faculty of Medicine University of Helsinki, Helsinki, Finland
9Translational Biology, Research and Development, Biogen Inc, Cambridge, Massachusetts, USA
10The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

Acknowledgements We want to acknowledge the participants and investigators of FinnGen study. The following biobanks are acknowledged for delivering biobank samples to FinnGen: Auria Biobank (www.auria.fi/biopankki), THL Biobank (www.thl.fi/biobank), Helsinki Biobank (www.helsinginbiopankki.fi), Biobank Borealis of Northern Finland (https://www.psph.fi/Tutkimus-ja-opetus/Biopankki/Pages/Biobank-Borealis-briefly-in-English.aspx), Finnish Clinical Biobank Tampere (www.tays.ensu.fi/Research_development/Finnish_Clinical_Biobank_Tampere), Biobank of Eastern Finland (www.itu-suomenbiopankki.fi/en), Central Finland Biobank (www.kkshp.fi/fifi/PottiLaale Biopankki), Finnish Red Cross Blood Service Biobank (www.veripalvelu.fi/verenluovutus/biopankkitilointa) and Terveystalo Biobank (www.terveystalo.com/fi/Viisytietoa/Terveystalo-Biopankki/Biopankki/). All Finnish biobanks are members of BMBRI infrastructure (www.bbrmi.nl). The Finnish Biobank Cooperative (FBB) (https://fbb.fi) is the coordinator of BMBRI-ERIC operations in Finland. The Finnish biobank data can be accessed through the Fingenious services (https://site.fingenious.fi/en/) managed by FINBB.

Collaborators FinnGen: Adam Platt (Astra Zeneca, Cambridge, United Kingdom), Adnan Ziegelstein (University of Michigan, Ann Arbor, MI, United States), Adriana Guerrero-Elizalde (University of Miami, Miami, FL, United States), The Many Brains Project, Boston, Massachusetts, USA

This study has also a few limitations. First, we could not assess the reasons for low participation because non-participants could not be further recontacted. Second, because this study could only recontact FinnGen participants, it essentially enrolled from a pool of individuals with a Finnish ancestry and a higher percentage of women than men. Therefore, generalising findings to other populations shall be done with caution.

In conclusion, this pilot suggests that recontacting individuals who have consented to be part of a biobank study with the goal to collect a wide range of cognitive, behavioural and lifestyle information can be challenging and may result in lower-than-expected participation. We speculate that returning some tangible incentive and/or relevant health information to participants might improve participation rates. Future studies are warranted to test this hypothesis and other strategies to improve participation in such survey research studies. Nonetheless, we suggest that cognitive, behavioural and lifestyle information be collected, whenever possible, at enrolment rather than via a post hoc recontacting process.
Data availability statement

Tampere (MH0004), Central Finland Biobank (1-2017) and Terveystalo Biobank (STB 2020_1), Finnish Clinical Biobank Tampere (HUS/359/2017), Auria Biobank (AB17-5154), Biobank Borealis of Northern Finland (BBO2020_1), Finnish Red Cross Blood Service Biobank (7.12.2017), Helsinki Biobank (VRK/6909/2018-3, VRK/4415/2019-3), the Social Insurance Institution (permit numbers: THL/2031/6.02.00/2017, THL/1101/5.05.00/2017, THL/341/6.02.00/2018, THL/2222/6.02.00/2018, THL/283/6.02.00/2019, THL/1721/5.05.00/2019, THL/1524/5.05.00/2020, THL/236/14.02/2020), Digital and Population Data Services Agency (permit numbers: VRK/3431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3), the Social Insurance Institution (permit numbers: KELA 522/2017, KELA 522/2018, KELA 522/2019, KELA 522/2020, KELA 138/522/2019, KELA 2/522/2020, KELA 16/522/2020); and the following industry partners: AbbVie, AstraZeneca UK, Biogen MA, Bristol Myers Squibb ( zahlung Corporation & Celgene International Ii Sarl), Genentech, Merck Sharp & Dohme, Pfizer, GlaxoSmithKline Intellectual Property Development, Sanofi US Services, Maze Therapeutics, Janssen Biotech, Novartis and Boehringer Ingelheim.

Competing interests

MEKN is a full-time employee at Novartis and HR is a full-time employee at Biogen.

Patient and public involvement

Patients and/or the public were not involved in the design, conduct, or reporting or dissemination plans of this research.

Patient consent for publication

Obtained.

Ethics approval

This study involves human participants. Patients and control subjects in FinnGen provided informed consent for biobank research, based on the Finnish Biobank Act. Alternatively, older research cohorts, collected prior the start of FinnGen (in August 2017), were collected based on study-specific consents and later transferred to the Finnish biobanks after approval by Fimea, the National Supervisory Authority for Welfare and Health. Recruitment protocols followed the biobank protocols approved by Fimea. The Coordinating Ethics Committee of the Hospital District of Helsinki and Uusimaa (HUS) approved the FinnGen study protocol (HUS/990/2017). The FinnGen study is approved by Finnish Institute for Health and Welfare (permit numbers: THL/2031/6.02.00/2017, THL/1101/5.05.00/2017, THL/341/6.02.00/2018, THL/2222/6.02.00/2018, THL/283/6.02.00/2019, THL/1721/5.05.00/2019, THL/1524/5.05.00/2020, THL/236/14.02/2020), Digital and Population Data Services Agency (permit numbers: VRK/3431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3), the Social Insurance Institution (permit numbers: KELA 522/2017, KELA 522/2018, KELA 522/2019, KELA 522/2020, KELA 138/522/2019, KELA 2/522/2020, KELA 16/522/2020), and Statistics Finland (permit numbers: TK-53-1041-17 and TK-53-90-20). The Biobank Access Decisions for FinnGen samples and data used in FinnGen Data Freeze 6 include: THL Biobank (BB2017_55, BB2017_111, BB2018_19, BB_2018_34, BB_2018_67, BB2018_71, BB2019_7, BB2019_8, BB2019_26, BB2020_1), Finnish Red Cross Blood Service Biobank (TRB 7.12.2017), Helsinki Biobank (HUS/359/2017), Auria Biobank (AB17-5154), Biobank Borrowers of Northern Finland (17_1013), Biobank of Eastern Finland (1186/2018), Finnish Clinical Biobank Tampere (M00004), Central Finland Biobank (1-2017) and Terveyttalo Biobank (STB 2018001). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review

Not commissioned; externally peer reviewed.

Data availability statement

Data are available upon reasonable request.
Open access

Supplemental material  This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access  This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID ID
Rodosthenis S Rodosthenous http://orcid.org/0000-0002-2312-5530

REFERENCES
1 Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature 2018;562:203–9.
2 Gaziano JM, Conrado J, Brophy M, et al. Million veteran program: a mega-biobank to study genetic influences on health and disease. J Clin Epidemiol 2016;70:214–23.
3 Karlson EW, Boutin NT, Hoffnagle AG, et al. Building the partners healthcare Biobank at partners personalized medicine: informed consent, return of research results, recruitment lessons and operational considerations. J Pers Med 2016;6. doi:10.3390/jpm6010002. [Epub ahead of print: 14 01 2016].
4 Nagai A, Hirata M, Karnatani Y, et al. Overview of the Biobank Japan project: study design and profile. J Epidemiol 2017;27:52–8.
5 Al Kuwari H, Al Thani A, Al Marri A, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to functional links to intelligence. Nat Genet 2018;50:363–71.
6 Kurki M, Karjalainen J, Palta P. FinnGen: unique genetic insights and analysis of 454,787 UK Biobank participants. Nat Commun 2019;10:1593–9.
7 Boyle J, Berman L, Dayton J, et al. Physical measures and biomarker collection in health surveys: propensity to participate. Res Social Adm Pharm 2021;17:921–9.
8 Pyykkönen A-J, Isomaa B, Pesonen A-K, et al. Sleep duration and insulin resistance in individuals without type 2 diabetes: the PPP-Megabank project. BMC Med Res Methodol 2019;19:230.
9 Pirastu N, Cordioli M, Nandakumar P, et al. Genetic analyses identify widespread sex-differential participation bias. Nat Genet 2021;53:663–71.
10 Al Kuwari H, Al Thani A, Al Marri A, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to functional links to intelligence. Nat Genet 2018;50:363–71.
11 Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature 2018;562:203–9.
12 Gaziano JM, Conrado J, Brophy M, et al. Million veteran program: a mega-biobank to study genetic influences on health and disease. J Clin Epidemiol 2016;70:214–23.
13 Karlson EW, Boutin NT, Hoffnagle AG, et al. Building the partners healthcare Biobank at partners personalized medicine: informed consent, return of research results, recruitment lessons and operational considerations. J Pers Med 2016;6. doi:10.3390/jpm6010002. [Epub ahead of print: 14 01 2016].
14 Nagai A, Hirata M, Karnatani Y, et al. Overview of the Biobank Japan project: study design and profile. J Epidemiol 2017;27:52–8.
15 Al Kuwari H, Al Thani A, Al Marri A, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to functional links to intelligence. Nat Genet 2018;50:363–71.
16 Kurki M, Karjalainen J, Palta P. FinnGen: unique genetic insights and analysis of 454,787 UK Biobank participants. Nat Commun 2019;10:1593–9.
17 Boyle J, Berman L, Dayton J, et al. Physical measures and biomarker collection in health surveys: propensity to participate. Res Social Adm Pharm 2021;17:921–9.
18 Pyykkönen A-J, Isomaa B, Pesonen A-K, et al. Sleep duration and insulin resistance in individuals without type 2 diabetes: the PPP-Megabank project. BMC Med Res Methodol 2019;19:230.
19 Pirastu N, Cordioli M, Nandakumar P, et al. Genetic analyses identify widespread sex-differential participation bias. Nat Genet 2021;53:663–71.
20 Al Kuwari H, Al Thani A, Al Marri A, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to functional links to intelligence. Nat Genet 2018;50:363–71.
21 Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature 2018;562:203–9.
22 Gaziano JM, Conrado J, Brophy M, et al. Million veteran program: a mega-biobank to study genetic influences on health and disease. J Clin Epidemiol 2016;70:214–23.
23 Karlson EW, Boutin NT, Hoffnagle AG, et al. Building the partners healthcare Biobank at partners personalized medicine: informed consent, return of research results, recruitment lessons and operational considerations. J Pers Med 2016;6. doi:10.3390/jpm6010002. [Epub ahead of print: 14 01 2016].
24 Nagai A, Hirata M, Karnatani Y, et al. Overview of the Biobank Japan project: study design and profile. J Epidemiol 2017;27:52–8.
25 Al Kuwari H, Al Thani A, Al Marri A, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to functional links to intelligence. Nat Genet 2018;50:363–71.
26 Kurki M, Karjalainen J, Palta P. FinnGen: unique genetic insights and analysis of 454,787 UK Biobank participants. Nat Commun 2019;10:1593–9.
27 Boyle J, Berman L, Dayton J, et al. Physical measures and biomarker collection in health surveys: propensity to participate. Res Social Adm Pharm 2021;17:921–9.
28 Pyykkönen A-J, Isomaa B, Pesonen A-K, et al. Sleep duration and insulin resistance in individuals without type 2 diabetes: the PPP-Megabank project. BMC Med Res Methodol 2019;19:230.
29 Pirastu N, Cordioli M, Nandakumar P, et al. Genetic analyses identify widespread sex-differential participation bias. Nat Genet 2021;53:663–71.