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

FinnGen

2022-10

FinnGen 2022 , ' Recontacting biobank participants to collect lifestyle, behavioural and cognitive information via online questionnaires : lessons from a pilot study within FinnGen ' , BMJ Open , vol. 12 , no. 10 , 064695 . https://doi.org/10.1136/bmjopen-2022-064695

http://hdl.handle.net/10138/350483
https://doi.org/10.1136/bmjopen-2022-064695

BMJ Open

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

Rodosthenis 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,[6] 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

To cite: Rodosthenous RS, Niemi MEK, Kallio L, *et al*. Recontacting biobank participants to collect lifestyle, behavioural and cognitive information via online questionnaires: lessons from a pilot study within FinnGen. *BMJ Open* 2022;**12**:e064695. doi:10.1136/bmjopen-2022-064695

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-064695).

Received 13 May 2022
Accepted 13 September 2022

For numbered affiliations see end of article.

Correspondence to
Andrea Ganna;
andrea.ganna@helsinki.fi

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.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ A scalable online portal was developed and tested to securely collect cognitive, behavioural and lifestyle information by recontacting FinnGen participants.
⇒ A genetic-based selection strategy was used to re-contact participants.
⇒ The effectiveness of different invitation methods to boost participation was tested.
⇒ Recontacting individuals with the aim to collect a broad range of information online can be challenging, especially in older individuals.
⇒ The study was conducted on a homogeneous population and the findings may not be generalisable to other populations.

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 insights[13 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,[17–19] 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.[20] 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. Samples from consented individuals can be used across many research projects, if approved by the biobanks. Thus, individuals are not actively informed of their participation in FinnGen. Here we describe the results of a pilot study that aimed at recontacting FinnGen participants with the goal to collect cognitive, behavioural and lifestyle information via a custom-made online platform.

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.[21] This was motivated by previous work showing a strong association between a PS for cognitive performances and study participation.[19 22] 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, multiracial 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[23] 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[21] and intelligence[24] using a multitrait analysis of GWAS method as described elsewhere.[25] 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 (a=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 (a=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

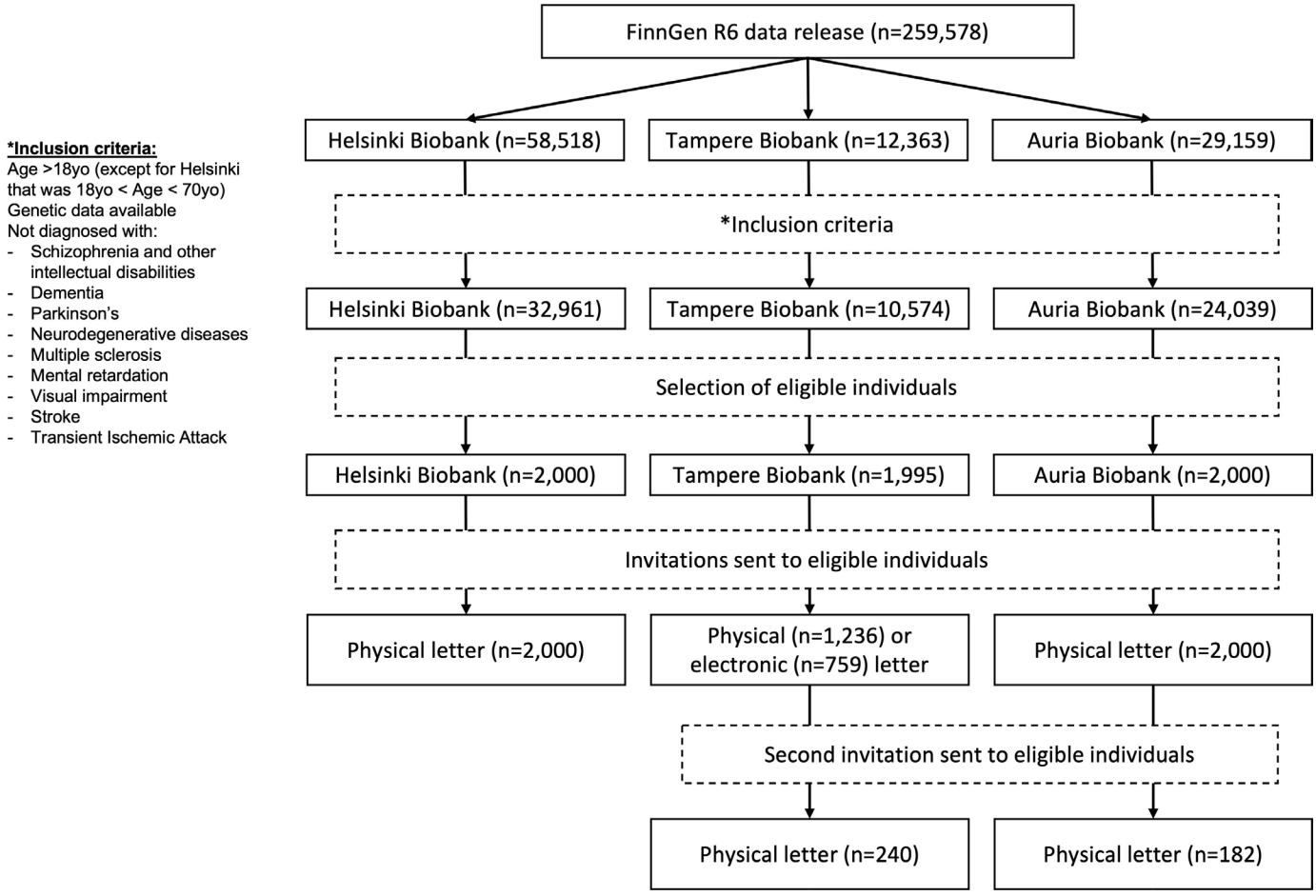

Figure 1 Flow diagram depicting the selection of study participants.

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).

Table 1 Basic questionnaire participation rates among all invitees by age group and biobank

Age	Helsinki (n=2000)			Tampere (n=1995)			Auria (n=2000)			All (n=5995)		
	Invited	Participated	Participation rate (%)	Invited	Participated	Participation rate (%)	Invited	Participated	Participation rate (%)	Invited	Participated	Participation rate (%)
20–29	149	22	14.8	116	18	15.5	41	8	19.5	306	48	15.7
30–39	426	70	16.4	273	51	18.7	115	23	20.0	814	144	17.7
40–49	379	93	24.5	231	44	19.0	159	49	30.8	769	186	24.2
50–59	484	133	27.5	303	58	19.1	218	64	29.4	1005	255	25.4
60–69	562	142	25.3	410	93	22.7	533	150	28.1	1505	385	25.6
70–79	0	0	0.0	503	62	12.3	686	21	3.1	1189	83	7.0
80–89	0	0	0.0	152	13	8.6	226	1	0.4	378	14	3.7
90–99	0	0	0.0	7	0	0.0	22	0	0.0	29	0	0.0
Total	2000	460	23	1995	339	17	2000	316	15.8	5995	1115	18.6

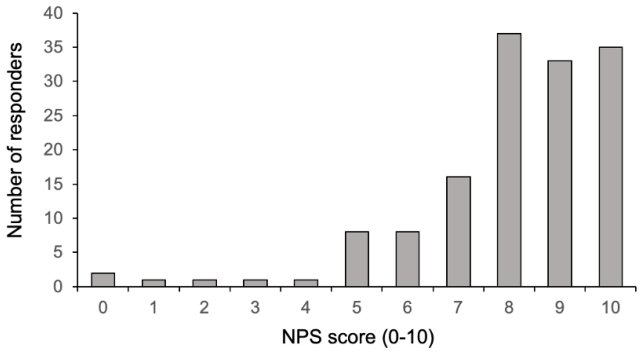

Figure 2 Net Promoter Scores (NPS) from the feedback survey asking for the experience of participants completing the online questionnaires (n=143).

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,

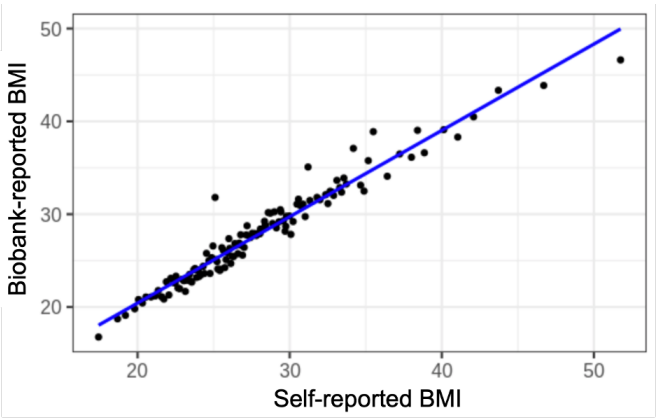

Figure 3 Spearman's correlation between biobank-reported and self-reported body mass index (BMI) measures acquired from the same participants (n=647).

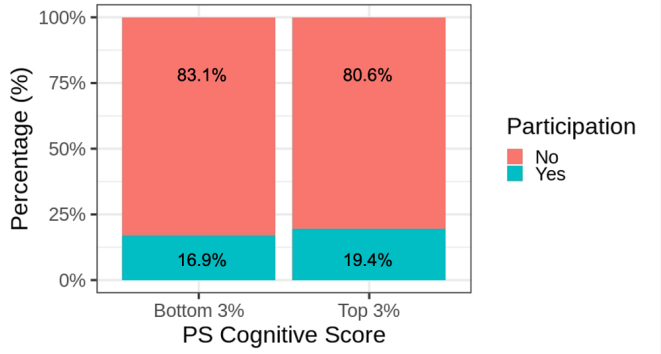

Figure 4 Stacked bar graph showing the participation proportion of study participants by polygenic score (PS) cognitive performances group.

respectively. Similarly, we could compare self-reported disease diagnoses from the questionnaire with disease information available in FinnGen and obtained from national health registers (online supplemental table 2). In the questionnaire, participants were asked whether they had been professionally diagnosed with any of the listed diseases. Overall, we noticed a higher prevalence for most diseases when self-reported compared with when extracted from the health registers. For example, asthma had a prevalence of 11.3% (124/1101) when data were extracted from the health registers compared with 18.5% (204/1101) when self-reported via the online questionnaire. We observed a similar trend with vitiligo: the prevalence of vitiligo from health register data was 0.1% (1/1101) vs 1.7% (19/1101) when self-reported. On the other hand, in the case of primary sclerosing cholangitis, we found the same prevalence of 0.7% (8/1101) between the health register and self-reported data. 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 diseases 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 (figure 4).

DISCUSSION

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,[26] research studies conducted in Finland have shown a high participation rate (>50%).[27–30] 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.[31 32]

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.[33] 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.[34] 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.[33]

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
[1]Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland
[2]Turku University Hospital (TYKS), Turku, Finland
[3]Hospital District of Helsinki and Uusimaa, Helsinki, Finland
[4]Tampere University Hospital, Tampere, Finland
[5]Finnish Institute for Health and Welfare, Helsinki, Finland
[6]Finnish Biobank Cooperative, Helsinki, Finland
[7]The Many Brains Project, Boston, Massachusetts, USA
[8]Department of Psychology and Logopedics, Faculty of Medicine University of Helsinki, Helsinki, Finland
[9]Translational Biology, Research and Development, Biogen Inc, Cambridge, Massachusetts, USA
[10]The 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.ppshp.fi/Tutkimus-ja-opetus/Biopankki/Pages/Biobank-Borealis-briefly-inEnglish.aspx), Finnish Clinical Biobank Tampere (www.tays.fi/enUS/Research_and_development/Finnish_Clinical_Biobank_Tampere), Biobank of Eastern Finland (www.ita-suomenbiopankki.fi/en), Central Finland Biobank (www.ksshp.fi/fiFI/Potilaalle/Biopankki), Finnish Red Cross Blood Service Biobank (www.veripalvelu.fi/verenluovutus/biopankkitoiminta) and Terveystalo Biobank (www.terveystalo.com/fi/Yritystietoa/Terveystalo-Biopankki/Biopankki/). All Finnish biobanks are members of BBMRI.fi infrastructure (www.bbmri.fi). The Finnish Biobank Cooperative (FINBB) (https://finbb.fi/) is the coordinator of BBMRI-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), Adam Ziemann (Abbvie, Chicago, IL, United States), Adriana Huertas-Vazquez (Merck, Kenilworth, NJ, United States), Aino Salminen (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Airi Jussila (Pirkanmaa Hospital District, Tampere, Finland), Aki Havulinna (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland; Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Alessandro Porello (Janssen Research & Development, LLC, Spring House, PA, United States), Ali Abbasi (Abbvie, Chicago, IL, United States), Amanda Elliott (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland; Broad Institute, Cambridge, MA, USA and Massachusetts General Hospital, Boston, MA, USA), Amy Hart (Janssen Research & Development, LLC, Spring House, PA, United States), Anastasia Kytölä (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Anders Mälarstig (Pfizer, New York, NY, United States), Andrey Loboda (Merck, Kenilworth, NJ, United States), Anne Lehtonen (Abbvie, Chicago, IL, United States), Anne Pitkäranta (Helsinki Biobank / Helsinki University and Hospital District of Helsinki and Uusimaa, Helsinki), Anne Remes (Northern Ostrobothnia Hospital District, Oulu, Finland), Annika Auranen (Pirkanmaa Hospital District, Tampere, Finland), Antti Aarnisalo (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Antti Hakanen (Auria Biobank / University of Turku / Hospital District of Southwest Finland, Turku, Finland), Antti Mäkitie (Department of Otorhinolaryngology - Head and Neck Surgery, University of Helsinki and Helsinki University Hospital, Helsinki, Finland), Antti Palomäki (Hospital District of Southwest Finland, Turku, Finland), Anu Jalanko (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Anu Loukola (Helsinki Biobank / Helsinki University and Hospital District of Helsinki and Uusimaa, Helsinki), Aoxing Liu (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Apinya Lertratanakul (Abbvie, Chicago, IL, United States), Argyro Bizaki-Vallaskangas (Pirkanmaa Hospital District, Tampere, Finland), Arto Lehisto (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Arto Mannermaa (Biobank of Eastern Finland / University of Eastern Finland / Northern Savo Hospital District, Kuopio, Finland), Athena Matakidou (Astra Zeneca, Cambridge, United Kingdom), Audrey Chu (GlaxoSmithKline, Brentford, United Kingdom), Auli Toivola (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland) Awaisa Ghazal (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Benjamin Challis (Astra Zeneca, Cambridge, United Kingdom), Bridget Riley-Gills (Abbvie, Chicago, IL, United States), Caroline Fox (Merck, Kenilworth, NJ, United States), Chia-Yen Chen (Biogen, Cambridge, MA, United States), Chris O'Donnell (Novartis Institutes for BioMedical Research, Cambridge, MA, United States) Clément Chatelain (Translational Sciences, Sanofi R&D, Framingham, MA, USA), Daniel Gordin (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), David Choy (Genentech, San Francisco, CA, United States), David Pulford (GlaxoSmithKline, Stevenage, United Kingdom), David Rice (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Dawn Waterworth (Janssen Research & Development, LLC, Spring House, PA, United States), Debby Ngo (Novartis, Basel, Switzerland), Deepak Raipal (Translational Sciences, Sanofi R&D, Framingham, MA, USA), Dermot Reilly (Janssen Research & Development, LLC, Boston, MA, United States), Diptee Kulkarni (GlaxoSmithKline, Brentford, United Kingdom), Dirk Paul (Astra Zeneca, Cambridge, United Kingdom), Edmond Teng (Genentech, San Francisco, CA, United States), Eeva Kangasniemi (Finnish Clinical Biobank Tampere / University of Tampere / Pirkanmaa Hospital District, Tampere, Finland), Eeva Sliz (University of Oulu, Oulu, Finland), Eija Laakkonen (University of Jyväskylä, Jyväskylä, Finland), Ekaterina Khramtsova (Janssen Research & Development, LLC, Spring House, PA, United States), Elina Järvensivu (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Elina Kilpeläinen (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Elisa Rahikkala (Northern Ostrobothnia Hospital District, Oulu, Finland), Elisabeth Widen (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Elmo Saarentaus (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Eric Green (Maze Therapeutics, San Francisco, CA, United States), Erich Strauss (Genentech, San Francisco, CA, United States), Erkki Isometsä (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Esa Pitkänen (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Essi Kaiharju (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Eveliina Salminen (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Fabiana Farias (Merck, Kenilworth, NJ, United States), Fanli Xu (GlaxoSmithKline, Brentford, United Kingdom), Fedik Rahimov (Abbvie, Chicago, IL, United States), Felix Vaura (Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Fredrik Åberg (Transplantation and Liver Surgery Clinic, Helsinki University Hospital, Helsinki University, Helsinki, Finland), George Okafo (Boehringer Ingelheim, Ingelheim am Rhein, Germany), Glenda Lassi (Astra Zeneca, Cambridge, United Kingdom), Hanna

Ollila (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Hannele Laivuori (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Hannele Mattsson (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Hannu Kankaanranta (University of Gothenburg, Gothenburg, Sweden/ Seinäjoki Central Hospital, Seinäjoki, Finland/ Tampere University, Tampere, Finland), Hannu Uusitalo (Pirkanmaa Hospital District, Tampere, Finland), Hao Chen (Genentech, San Francisco, CA, United States), Harri Siirtola (University of Tampere, Tampere, Finland), Heidi Silven (University of Oulu, Oulu, Finland), Heikki Joensuu (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Heli Lehtonen (Pfizer, New York, NY, United States), Heli Salminen-Mankonen (Boehringer Ingelheim, Ingelheim am Rhein, Germany), Henna Palin (Finnish Clinical Biobank Tampere / University of Tampere / Pirkanmaa Hospital District, Tampere, Finland), Henrike Heyne (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Hilkka Soininen (Northern Savo Hospital District, Kuopio, Finland), Howard Jacob (Abbvie, Chicago, IL, United States), Hubert Chen (Genentech, San Francisco, CA, United States), Iida Vähätalo (University of Tampere, Tampere, Finland), Iiris Hovatta (University of Helsinki, Finland), Ilkka Kalliala (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Ioanna Tachmazidou (Astra Zeneca, Cambridge, United Kingdom), Jaakko Kaprio (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Jaakko Parkkinen (Pfizer, New York, NY, United States), Jaakko Tyrmi (University of Oulu, Oulu, Finland / University of Tampere, Tampere, Finland), Jaana Suvisaari (Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Jae-Hoon Sul (Merck, Kenilworth, NJ, United States), Janet Kumar (GlaxoSmithKline, Collegeville, PA, United States), Jari Laukkanen (Central Finland Health Care District, Jyväskylä, Finland), Jarmo Ritari (Finnish Red Cross Blood Service, Helsinki, Finland), Jason Miller (Merck, Kenilworth, NJ, United States), Javier Garcia-Tabuenca (University of Tampere, Tampere, Finland), Jeffrey Waring (Abbvie, Chicago, IL, United States), Jenni Aittokallio (Hospital District of Southwest Finland, Turku, Finland), Jennifer Schutzman (Genentech, San Francisco, CA, United States), Jiwoo Lee (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland; Broad Institute, Cambridge, MA, United States), Joanna Betts (GlaxoSmithKline, Brentford, United Kingdom), Joel Rämö (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Johanna Huhtakangas (Northern Ostrobothnia Hospital District, Oulu, Finland), Johanna Mattson (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Johanna Schleutker (Auria Biobank / Univ. of Turku / Hospital District of Southwest Finland, Turku, Finland), Johannes Kettunen (Northern Finland Biobank Borealis / University of Oulu / Northern Ostrobothnia Hospital District, Oulu, Finland), John Eicher (GlaxoSmithKline, Brentford, United Kingdom), Joni A Turunen (Helsinki University Hospital and University of Helsinki, Helsinki, Finland; Eye Genetics Group, Folkhälsan Research Center, Helsinki, Finland), Jorge Esparza Gordillo (GlaxoSmithKline, Brentford, United Kingdom), Joseph Maranville (Bristol Myers Squibb, New York, NY, United States), Juha Karjalainen (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Juha Mehtonen (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Juha Rinne (Hospital District of Southwest Finland, Turku, Finland), Juha Sinisalo (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Jukka Koskela (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Jukka Partanen (Finnish Red Cross Blood Service / Finnish Hematology Registry and Clinical Biobank, Helsinki, Finland), Jukka Peltola (Pirkanmaa Hospital District, Tampere, Finland), Julie Hunkapiller (Genentech, San Francisco, CA, United States), Jussi Hernesniemi (Pirkanmaa Hospital District, Tampere, Finland), Juulia Partanen (Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland), Jyrki Pitkänen (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Kai Kaarniranta (Northern Savo Hospital District, Kuopio, Finland), Kaisa Tasanen (Northern Ostrobothnia Hospital District, Oulu, Finland), Kaj Metsärinne (Hospital District of Southwest Finland, Turku, Finland), Kalle Pärn (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Karen He (Janssen Research & Development, LLC, Spring House, PA, United States), Kari Eklund (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Katariina Hannula-Jouppi (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Katherine Klinger (Translational Sciences, Sanofi R&D, Framingham, MA, USA), Kati Donner (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Kati Hyvärinen (Finnish Red Cross Blood Service, Helsinki, Finland), Kati Kristiansson (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Katja Kivinen (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Katri Kaukinen (Pirkanmaa Hospital District, Tampere, Finland), Katri Pylkäs (University of Oulu, Oulu, Finland), Katriina Aalto-Setälä (Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland), Kimmo Palin (University of Helsinki, Helsinki, Finland), Kirsi Auro (GlaxoSmithKline, Espoo, Finland), Kirsi Kalpala (Pfizer, New York, NY, United States), Kirsi Sipilä (Research Unit of Oral Health Sciences Faculty of Medicine, University of Oulu, Oulu, Finland; Medical Research Center, Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland), Klaus Elenius (Hospital District of Southwest Finland, Turku, Finland), Kristiina Aittomäki (Department of Medical Genetics, Helsinki University Central Hospital, Helsinki, Finland), Kristin Tsuo (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland; Broad Institute, Cambridge, MA, United States), L. Elisa Lahtela (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Laura Addis (GlaxoSmithKline, Brentford, United Kingdom), Laura Huilaja (Northern Ostrobothnia Hospital District, Oulu, Finland), Laura Kotaniemi-Talonen (Pirkanmaa Hospital District, Tampere, Finland), Laura Pirilä (Hospital District of Southwest Finland, Turku, Finland), Laure Morin-Papunen (Northern Ostrobothnia Hospital District, Oulu, Finland), Lauri Aaltonen (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Leena Koulu (Hospital District of Southwest Finland, Turku, Finland), Liisa Suominen (Northern Savo Hospital District, Kuopio, Finland), Linda McCarthy (GlaxoSmithKline, Brentford, United Kingdom), Lotta Männikkö (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Ma'en Obeidat (Novartis Institutes for BioMedical Research, Cambridge, MA, United States), Maarit Niinimäki (Northern Ostrobothnia Hospital District, Oulu, Finland), Majd Mouded (Novartis, Basel, Switzerland), Malla-Maria Linna (Helsinki Biobank / Helsinki University and Hospital District of Helsinki and Uusimaa, Helsinki), Manuel Rivas (University of Stanford, Stanford, CA, United States), Marc Jung (Boehringer Ingelheim, Ingelheim am Rhein, Germany), Margaret G. Ehm (GlaxoSmithKline, Collegeville, PA, United States), Margit Pelkonen (Northern Savo Hospital District, Kuopio, Finland), Mari Kaunisto (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Maria Siponen (Northern Savo Hospital District, Kuopio, Finland), Marianna Niemi (University of Tampere, Tampere, Finland), Marja Vääräsmäki (Northern Ostrobothnia Hospital District, Oulu, Finland), Marja-Riitta Taskinen (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Mark Daly (Institute for Molecular Medicine, Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland; Broad Institute of MIT and Harvard; Massachusetts General Hospital), Mark McCarthy (Genentech, San Francisco, CA, United States), Markku Laukkanen (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Markku Voutilainen (Hospital District of Southwest Finland, Turku, Finland), Marla Hochfeld (Bristol Myers Squibb, New York, NY, United States), Mart Kals (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Martti Färkkilä (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Mary Pat Reeve (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Masahiro Kanai (Broad Institute, Cambridge, MA, United States), Meijian Guan (Janssen Research & Development, LLC, Spring House, PA, United States), Melissa Miller (Pfizer, New York, NY, United States), Mengzhen Liu (Abbvie, Chicago, IL, United States), Mervi Aavikko (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Mika Helminen (University of Tampere, Tampere, Finland), Mika Kähönen (Pirkanmaa Hospital District, Tampere, Finland), Mike Mendelson (Novartis, Boston, MA, United States), Mikko Arvas (Finnish Red Cross Blood Service / Finnish Hematology Registry and Clinical Biobank, Helsinki, Finland), Mikko Hiltunen (University of Eastern Finland, Kuopio, Finland), Mikko Kiviniemi (Northern Savo Hospital District, Kuopio, Finland), Minna Brunfeldt (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Minna Karjalainen (University of Oulu, Oulu, Finland), Minna Raivio (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Mitja Kurki (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland; Broad Institute, Cambridge, MA, United States), Mutaamba Maasha (Broad Institute, Cambridge, MA, United States), Nan Bing (Pfizer, New York, NY, United States), Natalia Pujol (Estonian biobank, Tartu, Estonia), Natalie Bowers (Genentech, San Francisco, CA, United States), Nathan Lawless (Boehringer Ingelheim, Ingelheim am Rhein, Germany), Neha Raghavan (Merck, Kenilworth, NJ, United States), Nicole Renaud (Novartis Institutes for BioMedical Research, Cambridge, MA, United States), Niko Välimäki (University of Helsinki, Helsinki, Finland), Nina Mars (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Nina Pitkänen (Auria Biobank / University of Turku / Hospital District of Southwest Finland, Turku, Finland), Nizar Smaoui (Abbvie, Chicago, IL, United States), Oili Kaipiainen-Seppänen (Northern Savo Hospital District, Kuopio, Finland), Olli Carpén (Helsinki Biobank / Helsinki University and Hospital District of Helsinki and Uusimaa, Helsinki), Oluwaseun Alexander Dada (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Oskari Heikinheimo (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Outi Tuovila (Business Finland, Helsinki, Finland), Outi Uimari (Northern Ostrobothnia Hospital District, Oulu, Finland), Päivi Auvinen (Northern Savo Hospital District, Kuopio, Finland), Päivi Laiho (THL Biobank / Finnish Institute for Health and Welfare

(THL), Helsinki, Finland), Päivi Mäntylä (Northern Savo Hospital District, Kuopio, Finland), Paula Kauppi (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Peeter Karihtala (Northern Ostrobothnia Hospital District, Oulu, Finland), Pekka Nieminen (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Pentti Tienari (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Petri Virolainen (Auria Biobank / University of Turku / Hospital District of Southwest Finland, Turku, Finland), Pia Isomäki (Pirkanmaa Hospital District, Tampere, Finland), Pietro Della Briotta Parolo (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Pirkko Pussinen (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Priit Palta (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Qingqin S Li (Janssen Research & Development, LLC, Titusville, NJ 08560, United States), Raimo Pakkanen (Business Finland, Helsinki, Finland), Raisa Serpi (Northern Finland Biobank Borealis / University of Oulu / Northern Ostrobothnia Hospital District, Oulu, Finland), Rajashree Mishra (GlaxoSmithKline, Brentford, United Kingdom), Reetta Hinttala (Northern Finland Biobank Borealis / University of Oulu / Northern Ostrobothnia Hospital District, Oulu, Finland), Reetta Kälviäinen (Northern Savo Hospital District, Kuopio, Finland), Regis Wong (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Relja Popovic (Abbvie, Chicago, IL, United States), Rigbe Weldatsadik (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Riikka Arffman (University of Oulu, Oulu, Finland), Riitta Lahesmaa (Hospital District of Southwest Finland, Turku, Finland), Rion Pendergrass (Genentech, San Francisco, CA, United States), Risto Kajanne (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Robert Graham (Maze Therapeutics, San Francisco, CA, United States), Robert Plenge (Bristol Myers Squibb, New York, NY, United States), Robert Yang (Janssen Biotech, Beerse, Belgium), Roosa Kallionpää (Hospital District of Southwest Finland, Turku, Finland), Sahar Mozaffari (Maze Therapeutics, San Francisco, CA, United States), Sally John (Biogen, Cambridge, MA, United States), Sami Heikkinen (University of Eastern Finland, Kuopio, Finland), Sami Koskelainen (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Sampsa Pikkarainen (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Samuel Lessard (Translational Sciences, Sanofi R&D, Framingham, MA, USA), Samuli Ripatti (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Sanna Siltanen (Finnish Clinical Biobank Tampere / University of Tampere / Pirkanmaa Hospital District, Tampere, Finland), Sanna Toppila-Salmi (University of Helsinki, Finland), Sanni Lahdenperä (Biogen, Cambridge, MA, United States), Sanni Ruotsalainen (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Sarah Smith (Finnish Biobank Cooperative – FINBB), Satu Strausz (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Sauli Vuoti (Janssen-Cilag Oy, Espoo, Finland), Shabbeer Hassan (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Shameek Biswas (Bristol Myers Squibb, New York, NY, United States), Shanmukha Sampath Padmanabhuni (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Shuang Luo (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Simonne Longerich (Merck, Kenilworth, NJ, United States), Sini Lähteenmäki (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Sirkku Peltonen (Hospital District of Southwest Finland, Turku, Finland), Sirpa Soini (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Stefan McDonough (Pfizer, New York, NY, United States), Stephanie Loomis (Biogen, Cambridge, MA, United States), Susan Eaton (Biogen, Cambridge, MA, United States), Susanna Lemmelä (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Susanna Savukoski (University of Oulu, Oulu, Finland), Taneli Raivio (Helsinki Biobank / Helsinki University and Hospital District of Helsinki and Uusimaa, Helsinki), Tarja Laitinen (Finnish Clinical Biobank Tampere / University of Tampere / Pirkanmaa Hospital District, Tampere, Finland), Taru Tukiainen (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Teea Salmi (Pirkanmaa Hospital District, Tampere, Finland), Teemu Niiranen (Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Teemu Paajanen (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Teijo Kuopio (Central Finland Biobank / University of Jyväskylä / Central Finland Health Care District, Jyväskylä, Finland), Terhi Kilpi (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Terhi Ollila (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Terhi Piltonen (Northern Ostrobothnia Hospital District, Oulu, Finland), Terttu Harju (Northern Ostrobothnia Hospital District, Oulu, Finland), Thomas Damm Als (Aarhus University, Denmark), Tiina Luukkaala (University of Tampere, Tampere, Finland), Tiinamaija Tuomi (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Tim Lu (Genentech, San Francisco, CA, United States), Timo Blomster (Northern Ostrobothnia Hospital District, Oulu, Finland), Timo Hiltunen (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Timo P. Sipilä (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Tom Southerington (Finnish Biobank Cooperative – FINBB), Tomi P. Mäkelä (HiLIFE, University of Helsinki, Finland, Finland), Triin Laisk (Estonian biobank, Tartu, Estonia), Tuomo Kiiskinen (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Tuomo Mantere (Northern Finland Biobank Borealis / University of Oulu / Northern Ostrobothnia Hospital District, Oulu, Finland), Tuomo Meretoja (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Tuula Palotie (University of Helsinki and Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Tuula Salo (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Tuuli Sistonen (THL Biobank / Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Tytti Willberg (Hospital District of Southwest Finland, Turku, Finland), Ulla Palotie (Hospital District of Helsinki and Uusimaa, Helsinki, Finland), Ulvi Gursoy (Hospital District of Southwest Finland, Turku, Finland), Valtteri Julkunen (Northern Savo Hospital District, Kuopio, Finland), Varpu Jokimaa (Hospital District of Southwest Finland, Turku, Finland), Veikko Salomaa (Finnish Institute for Health and Welfare (THL), Helsinki, Finland), Veli-Matti Kosma (Biobank of Eastern Finland / University of Eastern Finland / Northern Savo Hospital District, Kuopio, Finland), Veli-Matti Kosma (Biobank of Eastern Finland / University of Eastern Finland / Northern Savo Hospital District, Kuopio, Finland), Venla Kurra (Pirkanmaa Hospital District, Tampere, Finland), Vincent Llorens (Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland), Vuokko Anttonen (Northern Ostrobothnia Hospital District, Oulu, Finland), Wei Zhou (Broad Institute, Cambridge, MA, United States), Xinli Hu (Pfizer, New York, NY, United States), Ying Wu (Pfizer, New York, NY, United States), Zhihao Ding (Boehringer Ingelheim, Ingelheim am Rhein, Germany).

Contributors Conceptualisation: AG, MEKN, RSR, MP, AP, JL. Methodology: AG, MEKN, RSR, PW, MH, CM, PM, LG, VAM, OAK. Project administration and investigation: RSR, LK, MP, PT, TK, EP, EMM, PN, JM, MP, HS. Data curation: MEKN, RSR, PT, PN, TK, RSR, TSH, TJ. Formal analysis: RSR, MEKN, AG. Writing original draft: RSR, AG. Writing, review and editing: JL, HR, AP, MP, AG. Guarantor: AG.

Funding The FinnGen project is funded by two grants from Business Finland (HUS 4685/31/2016 and UH 4386/31/2016) and the following industry partners: AbbVie, AstraZeneca UK, Biogen MA, Bristol Myers Squibb (and Celgene Corporation & Celgene International II Sàrl), 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, or 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/2364/14.02/2020), Digital and Population Data Services Agency (permit numbers: VRK43431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3), the Social Insurance Institution (permit numbers: KELA 58/522/2017, KELA 131/522/2018, KELA 70/522/2019, KELA 98/522/2019, 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 (7.12.2017), Helsinki Biobank (HUS/359/2017), Auria Biobank (AB17-5154), Biobank Borealis of Northern Finland (2017_1013), Biobank of Eastern Finland (1186/2018), Finnish Clinical Biobank Tampere (MH0004), Central Finland Biobank (1-2017) and Terveystalo 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.

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, Concato 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, Kamatani Y, *et al*. Overview of the Biobank Japan project: study design and profile. *J Epidemiol* 2017;27:S2–8.
5 Al Kuwari H, Al Thani A, Al Marri A, *et al*. The Qatar Biobank: background and methods. *BMC Public Health* 2015;15:1208.
6 Kuriyama S, Yaegashi N, Nagami F, *et al*. The Tohoku medical Megabank project: design and mission. *J Epidemiol* 2016;26:493–511.
7 Chen Z, Chen J, Collins R, *et al*. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol* 2011;40:1652–66.
8 Wei CY, Yang JH, Yeh EC. Genetic profiles of 103,106 individuals in the Taiwan Biobank provide insights into the health and history of Han Chinese. NPJ Genom Med 2021;6:10.
9 Finer S, Martin HC, Khan A, *et al*. Cohort Profile: East London Genes & Health (ELGH), a community-based population genomics and health study in British Bangladeshi and British Pakistani people. *Int J Epidemiol* 2020;49:20–1.
10 Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. *Nat Genet* 2018;50:1593–9.
11 Backman JD, Li AH, Marcketta A, *et al*. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature* 2021;599:628–34.
12 Kanai M, Akiyama M, Takahashi A, *et al*. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat Genet* 2018;50:390–400.
13 Ganna A, Ingelsson E. 5 year mortality predictors in 498,103 UK Biobank participants: a prospective population-based study. *Lancet* 2015;386:533–40.
14 Richmond RC, Anderson EL, Dashti HS, *et al*. Investigating causal relations between sleep traits and risk of breast cancer in women: Mendelian randomisation study. *BMJ* 2019;365:l2327.
15 Suhre K, McCarthy MI, Schwenk JM. Genetics meets proteomics: perspectives for large population-based studies. *Nat Rev Genet* 2021;22:19–37.
16 Sudlow C, Gallacher J, Allen N, *et al*. Uk Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;12:e1001779.
17 Swanson JM. The UK Biobank and selection bias. *Lancet* 2012;380:110.
18 Munafò MR, Tilling K, Taylor AE, *et al*. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol* 2018;47:226–35.
19 Pirastu N, Cordioli M, Nandakumar P, *et al*. Genetic analyses identify widespread sex-differential participation bias. *Nat Genet* 2021;53:663–71.
20 Kurki MI, Karjalainen J, Palta P. FinnGen: unique genetic insights from combining isolated population and national health register data. *medRxiv*:2022:2022.03.03.22271360.
21 Lee JJ, Wedow R, Okbay A, *et al*. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet* 2018;50:1112–21.
22 Tyrrell J, Zheng J, Beaumont R, *et al*. Genetic predictors of participation in optional components of UK Biobank. *Nat Commun* 2021;12:886.
23 Wechsler D. *WAIS-R: Wechsler adult intelligence scale-revised*. New York, NY: Harcourt Brace Jovanovich [for] Psychological Corp., ©1981, 1981: 1896–981.
24 Savage JE, Jansen PR, Stringer S, *et al*. Genome-Wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet* 2018;50:912–9.
25 Turley P, Walters RK, Maghzian O, *et al*. Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat Genet* 2018;50:229–37.
26 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.
27 Pyykkönen A-J, Isomaa B, Pesonen A-K, *et al*. Sleep duration and insulin resistance in individuals without type 2 diabetes: the PPP-Botnia study. *Ann Med* 2014;46:324–9.
28 Borodulin K, Tolonen H, Jousilahti P, *et al*. Cohort profile: the National FINRISK study. *Int J Epidemiol* 2018;47:696–696i.
29 Ruokolainen O, Ollila H, Karjalainen K. Determinants of electronic cigarette use among Finnish adults: results from a population-based survey. *Nordisk Alkohol Nark* 2017;34:471–80.
30 Lobier M, Niittymäki P, Nikiforow N, *et al*. FinDonor 10 000 study: a cohort to identify iron depletion and factors affecting it in Finnish blood donors. *Vox Sang* 2020;115:36–46.
31 Lindsted KD, Fraser GE, Steinkohl M, *et al*. Healthy volunteer effect in a cohort study: temporal resolution in the Adventist health study. *J Clin Epidemiol* 1996;49:783–90.
32 Pinsky PF, Miller A, Kramer BS, *et al*. Evidence of a healthy volunteer effect in the prostate, lung, colorectal, and ovarian cancer screening trial. *Am J Epidemiol* 2007;165:874–81.
33 Harrison S, Henderson J, Alderdice F, *et al*. Methods to increase response rates to a population-based maternity survey: a comparison of two pilot studies. *BMC Med Res Methodol* 2019;19:65.
34 Smith MG, Witte M, Rocha S, *et al*. Effectiveness of incentives and follow-up on increasing survey response rates and participation in field studies. *BMC Med Res Methodol* 2019;19:230.

Supplementary Table 1. Basic questionnaire participation rates among all invitees under 70 years old by age group and biobank.

	Helsinki (n=2000)			**Tampere (n=1333)**			**Auria (n=1066)**			**All (n=4399)**		
Age	Invited	Participated	Participation Rate (%)	Invited	Participated	Participation Rate (%)	Invited	Participated	Participation Rate (%)	Invited	Participated	Participation Rate (%)
20-29	149	22	14,8	116	18	15,5	41	8	19,5	306	48	15,7
30-39	426	70	16,4	273	51	18,7	115	23	20,0	814	144	17,7
40-49	379	93	24,5	231	44	19,0	159	49	30,8	769	186	24,2
50-59	484	133	27,5	303	58	19,1	218	64	29,4	1005	255	25,4
60-69	562	142	25,3	410	93	22,7	533	150	28,1	1505	385	25,6
Total	**2000**	**460**	**23**	**1333**	**264**	**19,8**	**1066**	**294**	**27,6**	**4399**	**1018**	**23,1**

Supplementary Table 2. Prevalence of the medical history attainment for the same participants (n=1101) between the Re-contacting pilot and FinnGen studies.

	Prevalence, n (%)	
Medical conditions[1]	**Re-contacting (n=1101)**	**FinnGen (n=1101)**
Hypertension	313 (28.4%)	186 (16.9%)
Arthrosis/Osteoarthritis	291 (26.4%)	176 (16%)
Dry eyes	241 (21.9%)	51 (4.6%)
Depression	236 (21.4%)	55 (5%)
Asthma	204 (18.5%)	124 (11.3%)
Irritable bowel syndrome	139 (12.6%)	36 (3.3%)
Fatty liver disease/liver fibrosis	65 (5.9%)	21 (1.9%)
Fibromyalgia	47 (4.3%)	9 (0.8%)
Breast fibroadenomas	23 (2.1%)	16 (1.5%)
Vitiligo	19 (1.7%)	1 (0.1%)
Coeliac disease	14 (1.3%)	9 (0.8%)
Alopecia areata	13 (1.2%)	3 (0.3%)
Primary sclerosing cholangitis	8 (0.7%)	8 (0.7%)
Skin coeliac disease	3 (0.3%)	0 (0%)
None of the above	325 (29.5%)	0 (0%)
Do not know	18 (1.6%)	0 (0%)
Prefer not to answer	2 (0.2%)	0 (0%)

[1]In the Re-contacting study, all medical condition data were attained by asking participants "Have you ever been professionally diagnosed with any of the following disorders?". In FinnGen, the medical condition data were attained using the national health register data.