University of Groningen

Values of trust and participation in scientific data repositories

Broekstra, Reinder

DOI:
[10.33612/diss.195073701](https://doi.org/10.33612/diss.195073701)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Broekstra, R. (2021). *Values of trust and participation in scientific data repositories*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. <https://doi.org/10.33612/diss.195073701>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

For all who participate and do not participate.

Publisher: University of Groningen, Groningen, the Netherlands

Printer: Print.com

Cover & chapter illustrations: Simon Dam (simondam.com)

Funding: Dutch Research Council (NWO),
program Responsible Innovation, file number 313-99-313

Paranymphs: Tesfa Habtewold & Jesse Huisman

© 2021, Reinder Broekstra

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system of any nature, or transmitted in any form or by any means, electronic, mechanical, now known or hereafter invented, including photocopying or recording, without prior written permission from the copyright owner.

university of
 groningen

Values of trust and participation in scientific data repositories

PhD thesis

to obtain the degree of PhD at the
 University of Groningen
 on the authority of the
 Rector Magnificus Prof. C. Wijmenga
 and in accordance with
 the decision by the College of Deans.

This thesis will be defended in public on
 Wednesday december 8, 2021 at 16.15 hours

by

Reinder Broekstra

*born on 8 February 1988
 in Leeuwarden*

Supervisors

Prof. dr. R.P. Stolk

Prof. dr. S. Otten

Dr. E.L.M. Maeckelberghe

Co-supervisors

Dr. J.L. Aris-Meijer

Assessment Committee

Prof. dr. A.L. Bredenoord

Prof. dr. R. Bekkers

Prof. dr. H. Snieder

Content

General Summary	8
------------------------	----------

Chapter 1 General Introduction	15
References	20

Chapter 2 Motives of contributing personal data for health research: participation/non-participation in a Dutch biobank	22
--	-----------

Abstract	22
Introduction	23
Methods	24
Results	29
Discussion	34
Conclusion	37
References	38

Chapter 3 Trust in Large-Scale, Centralized Data Repository: A Qualitative Analysis **44**

Abstract	44
Introduction	45
Method	47
Results	52
Discussion	61
References	68

Chapter 4 Demographic and prosocial intrapersonal characteristics of biobank participants and refusers: the findings of a survey in the Netherlands **76**

Abstract	76
Introduction	77
Materials and methods	78
Results	83
Discussion	88
References	91

Chapter 5 Motives for withdrawal of participation in biobanking and participants' willingness to allow linkages of their data

96

Abstract	96
Introduction	97
Materials and Methods	100
Results	103
Discussion	111
References	117

Chapter 6 Dutch cohort studies and social capital: an examination of trust and participation indicators

124

Abstract	124
Introduction	125
Methods	127
Results	134
Conclusions and discussion	146
References	150

Mixed motives for participation/non-participation	164
Trust as a predictor of participation/non-participation	165
Implications for recruitment and retainment of participants	166
Implications for governance of data repositories	170
Limitations	172
References	173

Algemene samenvatting	178
Kokkuvõte	186
Research Institute SHARE theses	189
Acknowledgments	192
About the author	194

General Summary

Over the past thirty years, data storage has shifted from storage in physical repositories to storage in digital repositories. This digitalization has fostered the generation, storage, use and exchange of data, which has led to digital data repositories becoming a key resource for scientific research. Nowadays, an increasing amount of personal data can easily be acquired via sensor technologies, data linkage or open sources such as social media. Taking into consideration the increase in the number of people and technological devices producing data, there are opportunities for new types of research using these digital data repositories. These types of research exceed limitations of the pre-digital period by searching for connections across vast datasets while relying on a variety of data repositories and research purposes.

Biomedical research is one of the research fields that is clearly affected by the new research opportunities, as the utilisation of big data has potential significance for the understanding and improvement of health and health-related behaviour on an individual level. These new opportunities require many individuals to contribute a wide variety of data. As a result, digital repositories that are used for research purposes become large-scale. Moreover, these repositories provide central access to different types of data in their repository. Illustrative examples of large-scale, centralized data repositories in biomedical research are biobanks and large cohort studies. These banks or studies centrally store large quantities of biological specimens as well as data for a long period of time extracted from repeated questionnaires and measurements in certain time frame. As a result, they hold a vast amount of sensitive data on a long term, such as health data or genetic data, which could impact a person's livelihood, quality of life and ability to participate in society.

There is a wide consensus about the disruptiveness that is associated with the transition to large-scale, centralized data repositories for research since it changes flows of – particularly sensitive – information and reframes ideas about the processes of research and the constitution of knowledge. The knowledge and information that are generated with these data repositories might provide a range of opportunities to benefit individuals and whole societies. At the same time, these opportunities could bring new threats to these individuals and societies. The understanding of health and health-related behaviour on an individual level for a longer period might have a far-reaching impact on clinical practice, fundamental rights and fundamental freedoms. It is therefore not surprising that participation in biobanks is not self-evident. While some people are happy to contribute their data to science or healthcare and in doing so, provide an in-

sight into their private sphere, others worry about the exploitation or safety of their data, for example, if they were to participate. Moreover, some participants change their minds about participation and withdraw. A relevant question is, therefore, how can we better understand these differences between participation/non-participation in research data repositories and evaluate their inevitable impact?

Previous studies indicate that different forms of trust in society, such as trust in other individuals or trust in healthcare institutes, might be key to understanding participation/non-participation. Therefore, the present dissertation further investigates the role of trust in participation/non-participation, thereby specifically focusing on societal trust (i.e. trust in governmental institutes and in other individuals), trust in research and trust in sharing personal data. In summary, the overarching aim of this dissertation is to contribute to a better understanding of participation/non-participation in data repositories for research. We use the context of medicine, in particular epidemiology, as the basis for our research by focusing on participation/non-participation in biobanks and large cohort studies as examples of large-scale, centralized data repositories comprising sensitive data. This basis enables us to subsequently explore the ethical, social and psychosocial understanding of the behavioural and moral mechanisms involved in participation/non-participation.

The principal research question of this dissertation is: *Which factors affect individuals' willingness to contribute their personal data to large-scale, centralized data repositories for scientific research, and in particular to biobanks?* We answered this question through two explorative qualitative studies about correlates of participation/non-participation in biobanking. More specifically, we carried out in-depth interviews with participants and former participants of a Dutch population-based biobank and with non-participants who did not contribute to this biobank. We subsequently conducted three quantitative studies to corroborate and further improve our understanding of the relevant mechanisms and determinants of participation/non-participation as identified in the interviews.

The first part of our research is described in Chapters 2 and 3. These chapters report on findings that are based on qualitative data. We explore and discuss the reasons why individuals do or do not contribute to a Dutch population-based biobank as an example of a large-scale, centralized data repository. We investigated and compared the views of current participants and non-participants on contributing to large-scale, centralized research data repositories and investigated the reasons why former participants decided to end their participation. Moreover, we provide information about why and under what conditions individuals are willing to trust a Dutch population-based biobank with their person-

al data and which elements contribute to diminishing this trust.

The second part of our research is described in Chapters 4 to 6. These chapters report on quantitative data. Quantitative data of participants and those declining to participate in a Dutch population-based biobank were compared to start with. We acquired insights into their differences by assessing the demographic and prosocial intrapersonal characteristics of respondents who participated or declined to participate in the Lifelines biobank and databank. We subsequently analysed the motives that participants have for participation and potential withdrawal, and how these motives were related to various characteristics of the participants. Lastly, we focused on whether and how individuals' contribution of data to a Dutch population-based biobank may relate to trust in society and to other civic contributions (e.g. volunteering or voting behaviour). This study explores, therefore, the relationship between social capital and participation in public goods, using data from Statistics Netherlands linked with Lifelines data.

Overall, the findings in this dissertation signal that individuals' levels of trust and prosocial values can, at least to some extent, explain participation in a population-based biobank. More specifically, Chapter 2 suggests that the decision of both participants and non-participants regarding their contribution to the biobank was motivated by a desire to help other individuals. Whereas participants perceived benefits and limited risks, non-participants and former participants were concerned about individual or societal harm. Our analysis indicated that levels of societal trust, such as trust in government or other individuals, and in prosocial values, defined as concern for others, can partly account for these different perceptions of participants and non-participants.

Chapter 3 adds to these findings by showing that differences in levels of trust depend strongly on four individual beliefs. First, the belief that a large-scale data repository benefits the public. Second, the belief in the organizational aspect of data collection and storage, for example, the safe collection of data, the appropriate scale of the data repository, the sound reputation or trustworthiness of the responsible organization and the motives, behaviours and competence demonstrated by employees. Third, the belief that the data or insights, extracted from the collected data, are sensitive depending on range and context of repository. Fourth, the belief in the application of informed consent to retain control over personal data. These beliefs appeared to influence trust and ultimately the decision to participate or not and the decision to withdraw from participation.

Our quantitative studies support and extend the findings of our qualitative studies. Chapter 4 indicates that health-related values play a critical role in participation decisions besides levels of societal trust. The participation threshold

for Lifelines was determined by an absence of health-related values and of societal trust. Even though studies showed that motives for research participation are prosocial, we did not find a stronger general concern for others among participants compared to refusers of participation, when considering other values and trust levels. Nor did we find a higher trust in researchers among participants beyond the context of biobanking. Chapter 5 confirms that motives for participation and withdrawal were premised on benefits or harm to society and to the individuals themselves. The drivers of participants' willingness to accept data linkages for medical research include, in addition to a high level of trust in general, a strong motivation to realize individual and particularly societal benefits through participation and weak concerns about possible individual or societal harm. Chapter 6 tells us that if a person shows greater levels of commitment to society and trust in society, also known as social capital, the likelihood of them participating in both Lifelines and other data repositories for research is higher. While this commitment to society was reflected only in formal prosocial behaviours, such as voting or membership of non-profit organizations, trust in society reflected only a general trust in unknown others. In summary, the findings of this dissertation point to a dilemma in contributing to society without derogating individual interest and concerns. While participation in research data repositories for medical research requires participants to have strong prosocial values and orientations and a general trust in society (i.e. government and governmental researchers), the decision to participate relates both to appreciation of and experiences in the specific context of data repository and research (i.e. health and healthcare). Former participants provided insights into how their positive attitude towards participation changed due to a growing lack of societal trust or specifically trust in the biomedical research context. The research presented is the first to indicate how trust may affect the willingness to participate and, in particular, the readiness to withdraw from participation, depending on beliefs about society and on the appreciation of the context of data collection.

This dissertation underlines the variety of motives for participation in data repositories and participations' reciprocal character. Our findings show that the decision of participants and non-participants to contribute or not contribute to a biobank may, to some extent, be motivated by a desire to help others, particularly future generations. We found that the act of participation in data repositories for health research relates to prosocial attitudes and prosocial behaviours beyond the biomedical context, such as volunteering or participating in politics. This might indicate the importance of a common interest in participation. Yet, participation is not merely about helping others since our findings indicate that a lack of general prosocial orientations or other prosocial motives does not imply

non-participation. Health-related values were independently associated with participation in data repositories, which indicates the importance of individual interest. In addition to a common interest to do good and an individual interest in health and healthcare, two other factors relating to individual interest are likely to play an independent role in deciding on participation. These factors are individuals' research participation experience and experiences in their social network.

It might be possible to convince individuals to act prosocially and to participate by relating to motives other than helping others in society; for example, the motive to adhere to a norm in their social network or to fulfil an individual interest. By emphasizing the strengths of data repositories, both in terms of adding valuable information to information that is available in existing repositories and by providing individual benefits to the respondents (e.g. regular health check-ups), a wider variety of individuals could be recruited for participation. This could perhaps lead to an increasing number of people who have experience with participating in data repositories for research and to more participants from different social networks, thereby increasing participation in data repositories. This might also encourage groups that are currently left out in research.

Our data suggest that trust in society and trust in the relevant context of data repositories are strongly correlated but that both predict unique variance in participation/non-participation. Participants had more trust in government and in other individuals. At the same time, people who generally trust other individuals are more willing to be participants in multiple data repositories. However, participants' higher societal trust or trust in others did not imply a higher trust in commercial enterprises. Participants did not show stronger motives of societal harm in potential withdrawal decisions but they showed stronger motives of benefit and fewer concerns about individual harm. Altogether, these findings indicate how important positive beliefs about the result of research, trust in research and the trustworthiness of the research context are in deciding about participation in data repositories.

Hence, to increase the probability that individuals will participate in large-scale, centralized data repositories that apply these technologies, trust in the public and private sphere should be present to provide sufficient fuel for deciding in favour of participation. The opportunity as well as the challenge in safeguarding participation in data repositories for future innovative research is the increasing necessity of high levels of trust in general. It might, therefore, be key to inform individuals about the possibilities and limitations of data repositories from a broad perspective, providing information about the whole research context and its potential benefits and harm on an individual and a societal level.

This might improve potential participants' perceived control over data while enhancing trust, and might clarify which contextual norms apply and when.

Nowadays, biobanks and data repositories cross borders and contexts via complex systems, especially when applying data linkage and sensor data technologies. As a result, the responsibilities of large-scale, centralized data repositories might lie more on a system level, for example information and governance structures, than on an individual level, such as obtaining individual informed consent or establishing a secure data repository. The focus for data repositories in scientific research should, therefore, shift to recruitment systems for and governance of data that proactively enable informed decision-making about participation. This goes further than merely confirming that participants have sufficient knowledge about the risks and benefits in the respective research via an informed consent form or about acting on incidents such as data leakage. Organizations that are responsible for data repositories should ensure that governance systems and consequences of participation/non-participation in research are clearly and continuously communicated with participants and others in society, especially in population-based data repositories. Informed consent could then imply someone being informed about what they consent to and how that individual adapts their consent but also being informed about when and how you will be notified of situations that might require a renewal or adaptation of the consent given. A stepping stone could be increasing actual or potential participants' awareness of data impact by applying more visualizations of data processes by researchers themselves.

The responsibility to ensure that decisions on participation are informed and based on facts does not only lie in the hands of the organizations that set up data repositories. Governments and public institutes that represent research on a national and international level are key to the robust governance of research data repositories. They could and should help with informing, discussing and deciding about responsibilities and risks of potential societal and individual harm of data repositories. National and international governing bodies, such as data protection authorities or ethical review boards, should also further clarify opportunities and limitations for data repositories and strengthen proactive supervision of data repositories and their further use. Ensuring a better overview of the available data and databases might help to provide more guidance and clarity to the public and might foster the efficient use of data repositories. The establishment of data principles such as *FAIR* (*Findable, Accessible, Interoperable and Reusable*) or research infrastructures initiatives such as the *HEALTH-Research Infrastructure* could be important stepping stones towards an improvement of insight into and consensus of opportunities and limitations of data repositories.

Conclusion

In conclusion, values and trust provide rich grounds for understanding decisions regarding participation or non-participation, or withdrawal from data repositories for research. If participation decisions are made with expectations that are in line with reality and within thorough governance structures, these decisions honour a broad trust in society. It is not so much about the fact *that* a request for data is made but more about *how* this request is made and *what* the conditions surrounding the data request are. We hope that our findings will provide a relevant stepping stone for practical interventions directed at properly addressing potential participants' ethical and practical concerns and, subsequently, at enhancing and maintaining their trust and the probability that they will provide their personal data for scientific research. In this way, our present research, together with follow-up studies in this field, may contribute to the further spread of rich, large-scale data repositories on health-related and other societally relevant issues.

Chapter 1

General Introduction

Over the past thirty years, data storage has shifted from storage in physical repositories to storage in digital repositories. This digitalization has fostered the generation, storage, use and exchange of data, which has led to digital data repositories becoming a key resource for scientific research. Nowadays, an increasing amount of personal data can easily be acquired via sensor technologies, data linkage or open sources such as social media. Taking into consideration the increasing number of people and technological devices producing data, there are opportunities for new types of research using these digital data repositories. These types of research exceed limitations of the pre-digital period by searching for connections across vast datasets while relying on a variety of data repositories and research purposes [1].

Biomedical research is one of the research fields that is clearly affected by the new research opportunities, as the utilisation of big data has potential significance for the understanding and improvement of health and health-related behaviour on an individual level. For example, tailored medical treatment may become possible, based on a combination of genetic factors, environmental factors and psychosocial factors [2, 3]. This research requires a large number of individuals to contribute a wide variety of data. As a result, digital repositories that are used for research have become large-scale. Moreover, these repositories provide central access to different types of data in their repository. Illustrative examples of large-scale, centralized data repositories in biomedical research are biobanks and large cohort studies. Biobanks and large cohort studies have become key data repositories for the field of biomedicine, providing large amounts of personal data and biosamples (i.e. medical data) that are contributed by individuals in society for research purposes [4–8].

There is a wide consensus about the disruptiveness associated with the transition to large-scale, centralized data repositories since it changes information flows and reframes ideas about the research processes and the constitution of knowledge [9–11]. This development raises new ethical issues on a societal and an individual level on what to do and what not to do with those large-scale centralized data repositories [2]. Are conventional informed consent procedures still effective? What could happen to the privacy of individuals, both those who participate and those who do not? What could participants and non-participants gain or prevent by sharing their data, or by not sharing their data? All these developments and questions combined might have increased uncertainties about the impact of sharing personal data and conducting scientific research.

Data repositories are not a new development since the field of epidemiology already used large-scale data repositories as early as 1662. John Graunt analysed

counts of burials in London parish cemeteries and compiled thousands of death records in a single matrix – at that time clearly a large-scale data repository – that he used to investigate the causes of mortality [12]. A closer look reveals, however, that the potential impact of knowledge about individuals has strongly increased, as did the number of individuals and the volume of data [2]. While previously, detailed information about large numbers of individuals was difficult to acquire, such information is now available in large-scale, centralized data repositories, such as biobanks and large cohort studies. A very apt example to illustrate this change is that John Graunt limited his data collection to London while data from the World Health Organization about the COVID-19 pandemic have been obtained from all populated continents. Moreover, large-scale, centralized data repositories combine different types of data, such as behavioural data, demographic data or even financial data. This combination gives in-depth insight into a broad variety of individual characteristics and behaviours over time, which might all be related to individuals' health.

Although biobanks and large cohort studies can vary a lot, for example in their research objectives, number of participants, the amount and variety of data and the duration and interval of data collection, they inevitably contain a vast amount of personal data that are provided by a large number of individuals. Although still far from complete, Figure 1 shows that Europe, given its substantial and still increasing number of biobanks, may be on its way to creating an infrastructure that is representative of a substantial part of the European population [13].

Figure 1. Map of biobanks in 2015 listed in the directory of the Biobanks Biobanking and BioMolecular Resources Research Infrastructure-European Research Infrastructure Consortium [13].

The knowledge and information that are generated with health-related data repositories might provide a range of opportunities for individuals and for whole societies, by enabling researchers to generate knowledge about which individuals are at higher risk of developing cancer, for example. This could empower individuals to prevent clinical treatment by adjusting their behaviour. At the same time, it might help to improve clinical treatment and lower healthcare costs [12, 14]. Yet, these repositories could also bring new threats to individuals and to societies. Consider, for example, the possibility that knowledge on risk factors relating to cancer might also lead to the stigmatization, discrimination or exclusion of individuals who are at a higher risk of developing cancer from healthcare insurance [2]. An understanding of health and health-related behaviour on an individual level for a longer period of time might have a far-reaching impact on fundamental rights and freedoms, for example the right of privacy or the prohibition of discrimination [15]. Moreover, health-related data might provide sensitive information about direct participants and their relations, for example information about non-participating family members. It may relate directly to health (e.g. blood pressure) or indirectly (e.g. being a smoker). For these reasons, health-related data are considered to be among the most sensitive types of personal data since they provide an insight into a person's most intimate, or private, sphere [15, 16]. Unsurprisingly, then, participation in biobanks is not self-evident. While some people are happy to contribute their data to science or healthcare by providing an insight into their personal characteristics and circumstances, and in particular those related to their health, others worry about the exploitation or safety of their data, for example, if they were to participate. Although the existence of biobanks and studies shows that a substantial part of the population participates in such data repositories, studies also indicate that a large proportion of those who are asked to participate are not willing to do so [17]. Moreover, some participants change their minds about participation and withdraw from it. Nonetheless, the knowledge that is generated using these repositories aims to draw conclusions and suggest implications that apply to all individuals of a certain population, regardless of their participation.

A relevant question is, therefore, how can we better understand these differences between participation and non-participation in research data repositories and evaluate their inevitable impact? Several empirical studies have already shown that the decision to participate in scientific research might be seen as a form of prosocial behaviour in a public goods dilemma [18–20]. This perception considers the choice to participate as the result of placing common interests above self-interests, even though the benefits of participation may depend on how many other individuals participate. The decision to participate might also

be approached as a collective action problem [21–23]. From this perspective, each individual benefits from participation in a biobank by collectively reducing long-term health risks. This aim, however, can be undermined by non-participants who choose to pursue individual short-term gains rather than pay an individual price in the form of filling in questionnaires, providing biosamples or increasing risks to privacy, for example.

Previous studies indicate that different forms of trust in society might be key to understanding participation/non-participation [17, 24–27]. The stronger the dilemma in weighing and choosing between individual and common benefits and risks, the more relevant trust will be as a predictor of individual compliance for collective action or prosocial behaviour [28]. Therefore, the present dissertation further investigates the role of trust in participation/non-participation, thereby focusing specifically on trust in society, in research and in sharing personal data.

In summary, the overarching aim of this dissertation is to contribute to a better understanding of participation/non-participation in data repositories for research. We use the context of medicine, especially epidemiology, as the basis for our research by focusing on participation/non-participation in biobanks and large health-related cohort studies as examples of large-scale, centralized data repositories comprising sensitive data. This basis enables us to subsequently explore the ethical, social and psychosocial understanding of the behavioural and moral mechanisms involved in participation/non-participation.

The principal research question of this dissertation is: *Which factors affect individuals' willingness to contribute their personal data to large-scale, centralized data repositories for scientific research, and in particular to biobanks?* We tried to answer this question through two explorative qualitative studies about correlates of participation/non-participation in biobanking. More specifically, we carried out in-depth interviews with participants and former participants of a Dutch population-based biobank and with non-participants who were asked to contribute to this biobank. We* subsequently conducted three quantitative studies to corroborate and further improve our understanding of the relevant mechanism and determinants of participation/non-participation as identified in the interviews. Chapters 2–6 report on the research questions, the setup and the outcomes of our research. These chapters are based on articles that have either already been published or that have been submitted for publication. As a result, the chapters can be read individually while there will be some overlap in the sections describing the theoretical background and the main research questions.

Chapter 2 reports on findings that are based on qualitative data. It explores

and discusses the reasons why individuals do or do not contribute to a Dutch population-based biobank, Lifelines, as an example of a large-scale, centralized data repository. We investigated and compared the views of current participants and non-participants on contributing to large-scale, centralized research data repositories with those participants who withdrew. We conducted 36 in-depth interviews with former and current participants in a biobank and with non-participants to identify key factors. This comparison could yield new insights into motives for participation and non-participation, in particular the behavioural change associated with withdrawal.

Chapter 3 uses the same in-depth interview data as Chapter 2. It provides information about why and under which conditions individuals are willing to trust Lifelines with their personal data and which elements may contribute to diminishing this trust.

Chapter 4, as well as Chapters 5 and 6, reports on quantitative data. It compares participants in Lifelines with non-participants and aims to acquire insights into their differences by assessing the demographic and prosocial intrapersonal characteristics of respondents who participated ($n = 2,615$) and those who declined to participate ($n = 404$) in the Lifelines biobank and databank.

Chapter 5 also focuses on participation in a biobank. It analyses individuals' motives for participation and potential withdrawal. We also investigated how these motives were related to various characteristics of the participants. These questions were explored using a sample of participants in the Dutch Lifelines biobank ($n = 2,615$) who had expressed their willingness to share their data for research purposes.

Chapter 6 focuses on whether and how individuals' contribution of data to Lifelines may relate to trust in society and other civic contributions (e.g. volunteering or organ donation). To this end, it explores the relationship between social capital and participation in public goods, using unique data from Statistics Netherlands linked with the longitudinal Lifelines panel. We argued that being a respondent in the Lifelines biobank may be considered an act of charity towards society and may imply a sense of connectedness towards members of society. Hence, we assumed that people with more social capital are also more likely to be a participant in Lifelines.

Finally, in **Chapter 7** we reflect on the results of the previous chapters while answering the principal question of this thesis, namely 'which factors affect individuals' willingness to contribute their personal data to large-scale, centralized data repositories for scientific research, and in particular to biobanks?'. Some suggestions for further research and practical implications will also be provided.

References

- [1] Floridi L. Big data and their epistemological challenge. *Philos Technol* 2012; 25: 435–437.
- [2] Mittelstadt BD, Floridi L. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Sci Eng Ethics* 2016; 22: 303–341.
- [3] O'Doherty KC, Christofides E, Yen J, et al. If you build it, they will come: unintended future uses of organised health data collections. *BMC Med Ethics* 2016; 17: 16.
- [4] Krokstad S, Langhammer A, Hveem K, et al. Cohort profile: The HUNT study, Norway. *Int J Epidemiol* 2013; 42: 968–977.
- [5] Scholtens S, Smidt N, Swertz MA, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. *Int J Epidemiol* 2015; 44: 1172–1180.
- [6] 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: 1–10.
- [7] UK Biobank Coordinating Centre. UK Biobank: Protocol for a large-scale prospective epidemiological resource UK Biobank Coordinating Centre Stockport. *Design* 2007; 06: 1–112.
- [8] van Staa T-P, Goldacre B, Buchan I, et al. Big health data: the need to earn public trust. *Br Med J* 2016; 363: i3636.
- [9] Boyd D, Crawford K. Critical Questions for Big Data. *Information, Commun Soc* 2012; 15: 662–679.
- [10] Barocas S, Nissenbaum H. Big Data 's End Run around Anonymity and Consent. In: Julia L, Stodden V, Stefan B, et al. (eds) *Privacy, Big Data and the Public Good*. Cambridge; New York: Cambridge University Press, pp. 44–75.
- [11] Mayer-Schönberger V, Cukier K. *Big data : a revolution that will transform how we live, work and think*. London SE: John Murray, (2013).
- [12] Mooney SJ, Westreich DJ, El-Sayed AM. Epidemiology in the era of big data. *Epidemiology* 2015; 26: 390–394.
- [13] Holub P, Swertz M, Reihns R, et al. BBMRI-ERIC Directory: 515 Biobanks with over 60 Million Biological Samples. *Biopreserv Biobank* 2016; 14: 559–562.
- [14] Hamburg MA, Collins FS, MA H, et al. The path to personalized medicine. *N Engl J Med* 2010; 363: 301–304.

- [15] Mostert M, Bredenoord AL, Biesart MCIH, et al. Big Data in medical research and EU data protection law: Challenges to the consent or anonymise approach. *Eur J Hum Genet* 2016; 24: 956–960.
- [16] Riso B, Tupasela A, Vears DF, et al. Ethical sharing of health data in online platforms – which values should be considered? *Life Sci Soc Policy* 2017; 13: 12.
- [17] Gaskell G, Gottweis H, Starkbaum J, et al. Publics and biobanks: Pan-European diversity and the challenge of responsible innovation. *Eur J Hum Genet* 2013; 21: 14–20.
- [18] Parks CD, Joireman J, Van Lange PAM. Cooperation, Trust, and Antagonism. *Psychol Sci Public Interes* 2013; 14: 119–165.
- [19] Bekkers R. Traditional and health-related philanthropy: The role of resources and personality. *Soc Psychol Q* 2006; 69: 349–366.
- [20] Merz EM, Van Den Hurk K, De Kort WLAM. Organ donation registration and decision-making among current blood donors in the Netherlands. *Prog Transplant* 2017; 27: 266–272.
- [21] Healy K. Altruism as an Organizational Problem: The Case of Organ Procurement. *Am Sociol Rev* 2004; 69: 387–404.
- [22] Healy K. Do presumed-consent laws raise organ procurement rates. *DePaul Law Rev* 2005; 55: 1017.
- [23] Ladin K, Wang R, Fleishman A, et al. Does Social Capital Explain Community-Level Differences in Organ Donor Designation? *Milbank Q* 2015; 93: 609–641.
- [24] Van Lange PAM. Generalized Trust. *Curr Dir Psychol Sci* 2015; 24: 71–76.
- [25] Milne R, Morley KI, Howard H, et al. Trust in genomic data sharing among members of the general public in the UK, USA, Canada and Australia. *Hum Genet* 2019; 138: 1237–1246.
- [26] Voigt TH, Holtz V, Niemiec E, et al. Willingness to donate genomic and other medical data: results from Germany. *Eur J Hum Genet* 2020; 28: 1000–1009.
- [27] Critchley C, Nicol D, Otłowski M, et al. Predicting intention to biobank: A national survey. *Eur J Public Health* 2012; 22: 139–144.

Chapter 2

Motives of contributing personal data for health research: participation/ non-participation in a Dutch biobank

Abstract

Large-scale, centralized data repositories are playing a critical and unprecedented role in fostering innovative health research, leading to new opportunities as well as dilemmas for the medical sciences. Uncovering the reasons as to why citizens do or do not contribute to such repositories, for example, to population-based biobanks, is therefore crucial. We investigated and compared the views of existing participants and non-participants on contributing to large-scale, centralized health research data repositories with those of ex-participants regarding the decision to end their participation. This comparison could yield new insights into motives of participation and non-participation, in particular the behavioral change of withdrawal. We conducted 36 in-depth interviews with ex-participants, participants, and non-participants of a three generation, population-based biobank in the Netherlands. The interviews focused on the respondents' decision making processes relating to their participation in a large-scale, centralized repository for health research data. The decision of participants and non-participants to contribute to the biobank was motivated by a desire to help others. Whereas participants perceived only benefits relating to their participation and were unconcerned about potential risks, non-participants and ex-participants raised concerns about the threat of large-scale, centralized public data repositories and public institutes, such as social exclusion or commercialization. Our analysis of ex-participants' perceptions suggests that intrapersonal characteristics, such as levels of trust in society, participation conceived as a social norm, and basic societal values account for differences between participants and non-participants. Our findings indicate the fluidity of motives centering on helping others in decisions to participate in large-scale, centralized health research data repositories. Efforts to improve participation should focus on enhancing the trustworthiness of such data repositories and developing layered strategies for communication with participants and with the public. Accordingly, personalized approaches for recruiting participants and transmitting information along with appropriate regulatory frameworks are required, which have important implications for current data management and informed consent procedures.

Keywords: Trust, Big data, Biorepositories/Biobanks, Cohort study; Decision making; Qualitative Methods; The Netherlands, Justice/Participant Selection/Inclusion/Recruitment

This chapter is previously published as: Broekstra, R., Maeckelberghe, E.L.M., Aris-Meijer, J.L., Otten S., Stolk R. Motives of contributing personal data for health research: (non-) participation in a Dutch biobank. BMC Med Ethics 21, 62 (2020). <https://doi.org/10.1186/s12910-020-00504-3>

Introduction

There is an increasing collection and transmission of personal data in society. Biobanks and other large-scale, centralized data repositories are a special form of data collection, storage, and use of large volumes of diverse personal data obtained from patients and citizens. Their aim is to contribute to the development of research activities in medicine [1–3]. They constitute valuable resources for healthcare professionals and facilitate innovations in epidemiological, genetic, and public health research [2, 4]. These data repositories facilitate the efficient performance of large-scale and continuous data analyses and enable researchers and clinicians to generate individual tailored clinical care [5–9]. Contributing personal data to biomedical research must be voluntary [9–14].

There is widespread consensus among European researchers that contributions of personal data by citizens foster scientific development and are therefore commendable [14]. Studies show that participants also share this consensus. They are motivated to do good deeds by helping researchers, patients, and doctors [2, 15–17]. A lot of citizens nonetheless choose not to contribute personal data for medical scientific research purposes, since percentages of unwilling citizens in representative samples range from 17 to 44% [18, 19]. This raises the question why some individuals do (participants) and others (non-participants) do not actively contribute their personal data to data repositories for medical scientific research, such as population-based biobanks.

The current literature mentions two types of risk that might explain the decision to participate or not in medical scientific research [2, 15]. The first type are the foreseeable risks associated with repositories, such as increasing loss of direct control over personal data, loss of privacy, or data breaches. The second type includes less foreseeable risks, for example the risk of unauthorised re-identification with potentially harmful consequences that is incurred as a result of linking different data repositories for the conduct of new kinds of analyses [2, 17, 20]. This second type of risk is, in contrast to the first, a risk that is rarely addressed in currently used consent forms yet exists [2].

The two types of risk offer however insufficient explanation why some participants, which we refer to as ex-participants, withdraw during the data collection process. Several studies have suggested that withdrawal or other

non-participation is motivated by perceptions of the consequences of helping others. For example, concerns regarding the commercialization of public research that derogates social benefits intended for others [21–23], risks of privacy violations following individual and group data sharing [2, 19, 22, 24–34], or potential discrimination [35]. More insight is needed in what factors drive this behavioural change. A comparison of the experiences of ex-participants of a biobank and nonparticipants with the views of participants could therefore yield insights into the reasons for participation, non-participation, and, in particular, withdrawal.

Our aim is to understand how motives, values, and expectations can differ in participation and non-participation. This understanding of underlying mechanisms could yield insights on variance in ethical understandings of biobank research, improvement of research data governance in biobanking, and effective strategies of participant recruitment in biomedical research. In this qualitative study of ex-participants, participants, and non-participants of a Dutch population-based biobank, we investigated and compared the views of participants on contributing data to the biobank with those of non-participants and ex-participants about why they chose not to participate and why they ended their participation respectively.

Methods

Data collection

We conducted in-depth interviews with individuals who were ex-participants, non-participants, or participants in Lifelines, a Dutch population-based biobank, to explore the motives underlying their participation and non participation in a centralized health research data system. Lifelines is a multidisciplinary, prospective population-based cohort study that entails a unique three-generational design for examining the health and health-related behaviours of individuals living in the Northern Netherlands. Lifelines encompasses a broad range of investigative procedures used to assess the biomedical, sociodemographic, behavioural, physical, and psychological factors that contribute to health and disease within the general population, with a special focus on multi-morbidity and complex genetics [36]. The recruitment of participants began in 2006 and continued up to 2013. Recent year, Lifelines launched an initiative introducing innovative big data technologies that include, for example, providing research participants with personalized feedback.

We applied a narrative interview approach with a tailored topic guide

that was partly derived from the DIPex methodology that allowed for a discussion, clarification, and verification of unanticipated themes [37, 38]. The topic guide, as shown in Table 1, was developed in a research project about research (non-)participation and big data consisting of the current study and a qualitative study about trust [35]. Six relevant themes relating to the motives, values, and expectations associated with participation and non-participation were identified in this topic guide, based on the state of knowledge evidenced in the scientific literature on (non-)participation in biobanks, public goods, trust, and data sharing: (1) becoming a (non-)participant; (2) objective aspects of participation (e.g., participation overall, tasks, and feedback); (3) subjective aspects of participation (e.g., expectations of accomplishment and feelings of identification); (4) definitions of and attitudes towards big data; (5) perceived benefits and threats associated with big data; and (6) decisions to provide personal data. The topics discussed in the interviews were derived from these themes.

Table 1. Interview topics and illustrative questions

Topic	Example questions
Association with biobank	How did you get involved?
Decision of (non-)participation	What were your considerations?
Reality	Was your experience as expected or not?
Future	Will you participate in Lifelines or in another centralised large-scale research project in the future?
Overall perception	How is participation in Lifelines perceived in general?
Tasks	How are the tasks related to participation in Lifelines perceived?
Feedback	What do you think of the feedback on your data you (would) receive?
Outcome hoped for	What motivates individuals to participate?
Benefits	What are the benefits and opportunities relating to participation?
Disadvantages	What are the risks and disadvantages relating to participation?
Definition	What are centralised large-scale data repositories?
Attitudes towards large-scale centralised data repositories	Are the public's attitudes towards collection, linking and use of this data generally positive or negative?
Benefits of large-scale centralised data repositories	What are the benefits of collecting, linking and using data?
Risks posed by large-scale centralised data repository	What risks are posed by the collection, linking and use of data?
Personal balance of benefits and threats	From your personal perspective, how are these considerations balanced?
Societal balance of benefits and threats	From a societal perspective, how are these considerations balanced?

All of the interviews were conducted by a member of the research team (RB), who is a trained and experienced interviewer. Each interview covered all of the topics included in the guide and lasted between 30 and 65 min. Data saturation was reached, when informational redundancy was achieved in the interviews [39]. Current participants of Lifelines were interviewed following their regular visits to the Lifelines facility. Most of the interviews were held at the University Medical Center Groningen. All four ex-participants and eight out of fifteen non-participants were interviewed at their homes on their request. All respondents provided verbal informed consent for their participation and the recording of the inter-

view. This consent is documented by recording on video or writing. Apart from one interviewee, who only consented to a written recording of the interview, the remaining interviewees consented to audio recordings of their interviews. All of the audio recordings were transcribed by an independent professional organization. These recordings were anonymized and stored at the highly secured servers of the University Medical Center Groningen. The access to these servers is only with permission of the researchers and excludes access for staff of Lifelines. The Medical Ethics Review Board of the University Medical Center Groningen exempted this study from the ethics review according to the Medical Research Involving Human Subjects Act of the Netherlands, as it did not involve collection of any health-related data.

Recruitment of respondents and sampling

A maximum variation sampling strategy was applied to recruit participants for the study [40]. Thus, we ensured that the sample of selected interviewees was heterogeneous [41, 42]. Between August and September 2016, we interviewed 36 individuals: 17 participants, four ex-participants, and 15 non-participants in the Lifelines biobank. We refer to these distinct groups throughout this paper as ‘interviewees’ while distinguishing ‘participants’, ‘ex-participants’, and ‘non-participants’. Table 2 shows the characteristics of all of the interviewees. Their average age was 45 years (ranging between 20 and 68 years) and 17 out of the 36 interviewees were male.

We recruited individuals partly from the Lifelines biobank. To ensure their privacy, Lifelines recruited participants and ex-participants of the biobank without the researchers being aware of their identity. Ex-participants who had terminated their participation in the few months prior to the commencement of our study were approached by phone, using a protocol created by the researchers (RB and JA) during the final stage of Lifelines’ established withdrawal procedure. Moreover, the Lifelines biobank also invited non-participating partners of participants who had contributed to the biobank. Ex-participants were the least willing to participate in our study. Although 90 people were invited, only four (4.44%) consented to participate. Reasons given for not participating in our study were ‘no interest’, ‘not available’, and ‘done with participating’. The sample of ex-participants was too small to enable generalization, but yielded the in-depth information that we were seeking. Non-participants were recruited through direct encounters with members of the public at the entrance of Groningen’s central public library and by personal invitation vis-à-vis various social networks as well as through online and offline platforms.

Table 2. Respondents' Characteristics

Respondent	Sex	Age	Occupational Field	Participation status
1	Female	65-70	Human resources	Participant
2	Male	40-45	Transportation	Non-participant
3	Female	60-65	Education	Non-participant
4	Male	30-35	Legal and social work	Participant
5	Male	50-55	Photography and architecture	Ex-participant
6	Male	55-60	Municipality	Non-participant
7	Female	35-40	Home maintenance	Ex-participant
8	Male	25-30	Warehouse maintenance	Participant
9	Female	55-60	Retail	Participant
10	Male	30-35	Entrepreneur	Participant
11	Male	20-25	Student	Participant
12	Female	50-55	Law	Participant
13	Male	25-30	Student	Non-participant
14	Male	65-70	Mechanical engineering	Participant
15	Male	55-60	Home maintenance	Non-participant
16	Female	25-30	Publisher	Non-participant
17	Female	45-50	Entrepreneurship	Non-participant
18	Female	65-70	Retired	Participant
19	Male	50-55	Management	Participant
20	Female	25-30	Energy policy	Non-participant
21	Female	35-40	Healthcare	Participant
22	Female	50-55	Healthcare	Participant
23	Female	25-30	Human resources	Non-participant
24	Female	45-50	Education	Participant
25	Male	40-45	Municipality	Participant
26	Female	65-70	Farming	Non-participant
27	Female	40-45	Librarian	Ex-participant
28	Female	50-55	Social work	Ex-participant
29	Female	35-40	Management	Participant
30	Male	35-40	Engineering	Non-participant
31	Male	55-60	Human resources	Non-participant
32	Female	30-35	Home maintenance	Non-participant
33	Male	40-45	Management	Participant
34	Female	60-65	Retired	Participant
35	Male	35-40	Business development	Non-participant
36	Male	45-50	Entrepreneurship	Non-participant

Data analysis

We analysed every phrase in each of the 36 interview transcripts within the context of the entire interview, and, where appropriate, a code pertaining to its content was generated or assigned. Codes could be applied multiple times within each transcript, and phrases could comprise multiple codes. Codes with related content were clustered within groups that were subsequently thematically categorized. An initial coding protocol containing a description of the codes, groups, and themes was developed based on a close reading of three transcripts. This coding protocol was evaluated through an iterative process, whereby two or three researchers cross-checked analyses of each of the transcripts. Once a consensus regarding the content of the codes, groups, and themes had been reached, the resulting coding protocol was used for the remaining transcripts. Five researchers coded the transcripts. Six random transcripts were coded by at least three researchers. This procedure ensured agreement among the researchers regarding the coding. Subsequently, one researcher coded all of the remaining transcripts.

The results of the interviews were categorized according to four themes. One of these themes, ‘perception of big data’, that emerged from our topic guide was selected because it was a focal theme during the interviews prior to the analytical phase. The other themes did not emerge directly from our topic guide but instead featured prominently in the interview data. This approach can be considered as valid, as the adaptive theory for qualitative research allows for the influence of theory on research [43]. Transcripts were primarily analysed using the computer-assisted qualitative data analysis package, Atlas TI, version 8, to retrace and evaluate quotes along with their codes, groups, and themes [44].

Results

We distinguished five themes relating to the motives, values, and expectations of ex-participants, participants, and non-participants in our dataset. The first three themes pertained to views and motives regarding the initial or current participation of ex-participants and participants. The last two themes touched on the ex-participants’ views and motives relating to their withdrawal and on nonparticipants’ views.

Contribution to the public good: science, healthcare, and society

Initially, most of the participants perceived their participation in Lifelines in terms of the donation of their time, information, and data to a public good, namely health. They believed that they were helping society by voluntarily

investing time and effort, which they could also do in other ways. They compared their contribution of personal data to Lifelines to a gift bequeathed to a charity or to organ donation. For example, one participant (P33) explained: 'My role is not to conduct research but to provide things [such as information, money, or data]'.

Ex-participants emphasized that the central motive for their initial participation in Lifelines was their desire to contribute to society. According to them, the purpose of research was to 'gain insights' and to 'create innovation in society'. Therefore, in their view, participation would facilitate societal progress. An ex-participant (EP28) explained this as follows: 'There was at that time no reason to say "no". It was about gaining collective insights into the world; you need people and you need a research group'.

This perception of contributing to societal progress was fostered through the awareness that Lifelines is a prospective cohort and a large-scale, centralized data repository. Ex-participants argued that the scale and length of the Lifelines data collection was unique and could thus provide new insights for medical sciences in the long term. Their initial expectations were that the results and implications of the research conducted using Lifelines data would primarily benefit future generations. These expectations were similar to those of the participants. Two participants explained this as follows:

I think I have little benefit, I don't know. I think it's for the generation after me, for which I'm participating in now... My parents participated in such events for us, and so I am participating for the next generation. (P22).

Research must be useful for society as a whole [so] that we ultimately flourish or become less ill. ... so that things get better; so that we leave it better than we found it. It is that idealistic idea so to speak. (P1).

For the participants, the primary motive for participation was to contribute to science. Moreover, all of the participants and ex-participants subscribed to the belief that participation would or could facilitate scientific progress and specifically the progress of the medical sciences. Participants as well as ex-participants explicitly stated that they highly valued scientific research in general. Participants were primarily interested in improving health and medical knowledge because they prioritized health as a profoundly important value. They perceived their donations of personal data as contributions to health-related research that was aimed at facilitating a better understanding of the development and course of

illnesses. One participant (P22) expressed the view that ‘the medical sciences can’t stagnate, so you must cooperate in order to help people to do research and make it better. That is what my aim is’.

Participation to help family members or friends

Similarly, a motive for participation mentioned by the interviewees, especially ex-participants, was ‘helping family or friends’, which was described as an investment of time and effort to benefit the current generation of family members or friends. They perceived this as an important motive influencing their decision to participate. An ex-participant (EP27) made the following statement:

As I recall, I participated because a family member participated too. It was probably my mother, but I am not sure... We spoke about the (Lifelines) research on health and diseases at that time. However, exactly how that went I don’t recall because of the passage of time. I started participating after that conversation.

Several current Lifelines participants who were recruited by their friends or family members considered participation in the programme to be almost self-evident, given the objective of Lifelines to follow the trajectories of three generations. A participant (P18) explained this as follows:

Well, we have two daughters; one daughter took part in it and so did her son. So, I thought, “then someone from our generation will have to participate too”, so that they will have someone of my generation and get three generations in total. My husband didn’t want to participate, so I joined.

Contributing to obtain individual benefits

The prospect of a ‘medical check-up’ was also a motivating factor for the contributions of ex-participants. Once every 5 years, participants visit a Lifelines location, where they provide blood samples and are subjected to physical measurements (e.g., those relating to anthropometry, electrocardiogram, and spirometry). Afterwards, some of these measurements are returned to the participants. Interviewees perceived this feedback as a free medical check-up that provided them with extra health-related information. These same results also sent to the general practitioner (GP) with a recommendation on whether the GP should actively contact the participant to arrange further clinical investigations.

Although participants highly appreciated this sharing of data, which they

valued as a supplementary bonus to their personal healthcare, most of them emphasized that it was not the primary motive for their participation. They valued some of the other individual benefits a well, such as ‘acquiring knowledge in general’, ‘learning about themselves’, and ‘the joy of gaining insights into the scientific process’. Despite the fact that all of the participants were informed that not all of the measured data would be returned and that the collected data would not be screened for diseases, they felt that the chances of obtaining an early diagnosis of illness would increase through their participation in the programme. They also assumed that they were free of any illness if their GPs did not recommend further investigations. One participant (P19) noted:

I do like it. Suppose I am not in good health; it is of course out of self-interest as well, to find out at the earliest possible stage. Look, normally if you don't have anything, then you don't go to a doctor. Nevertheless, here an EEG is done, your blood is measured, your urine is measured, and other things are done. So if there is something wrong. .. then hopefully you will know at an early stage. So, that is also one of the reasons for me [to participate].

A few participants thought that their own data could prove useful if they were later to fall ill. Participation would therefore enable them to keep records of their personal health. For example, according to one participant (P11), ‘it is a sort of back up. We [my GP and I] can look into the past, if I feel ill or have some symptoms. Perhaps I [my GP and I] can use it?’ Applying this line of reasoning, some of the participants specifically observed that their participation and procurement of data could provide their offspring with extra health related information.

Not contributing to a public good

Prior to their withdrawal from the programme, ex-participants considered the large-scale, centralized data repository to be a public good. However, this positive perception later changed to suspicion, as they felt that their personal data would be sold or misused. The new perception was driven by sceptical news items, negative perceptions of family members and friends, or a negative experience during a research visit. They mentioned that their concerns about the data repository being used by insurance companies or by the government to exclude people collectively and individually were among the primary reasons why they withdrew from the Lifelines programme. Another reason pertained to their belief that Lifelines was a profit-oriented organization and, therefore, their participation did not contribute to a public good.

Non-participants also considered the risk of collective exclusion with Lifelines to be plausible. For example, they noted the effect of personalized information on insurance premiums that could derogate a 'social society'. Some of the non-participants pointed to historical evidence of the isolation and exclusion of social groups resulting from the misuse of comparable repositories or systems. One non-participant (NP36) specifically referred to the civilian administrative system that prevailed in the Netherlands prior to and during World War II: '... they used to register the religions of civilians. That is not done any more for good reasons. These practices originate from World War II and the [treatment of the] Jews. Those [data] were used to trace them'.

Similar to ex-participants, non-participants generally thought that their participation in large-scale, centralized data repositories could prompt a decrease in life opportunities for themselves and others through decisions on, for example, mortgages or governmental support. One non-participant (NP6) elaborated on this point as follows:

Well look, if it [the data] is linked to healthcare insurance, they could potentially determine from this information that this male or female has a risk that is above average. Therefore, they may ask for a higher premium or deny the individual insurance. That is. .. I perceive it as a definite risk and threat if such organizations are involved.

Differing from ex-participants, however, non-participants ranked the type of research compiled for population-based biobanks lower than other types of research in terms of its importance. Several non-participants were not particularly concerned about health research, showing considerably more interest in topics relating to economics or energy related innovations. Although they saw some use for research facilitated by Lifelines data, they did not feel that contributing their personal data would be of much use, as a non-participant (NP20) explained:

I don't necessarily perceive it [as being] negative; nor do I immediately think: "oh yes, I am going to participate". I think that it depends very much on. .. as you explained it now, it [research participation] doesn't attract me. I don't think: "oh yes I have to participate, because I can be of use" or something like that.

A lack of individual benefits

Ex-participants explained that one of the reasons for their withdrawal was that they did not perceive their participation as being useful in terms of procuring personal benefits that matched their expectations. They discovered that participants could still become ill regardless of their participation

in Lifelines, indicating a lack of individual benefit of a ‘medical check-up’. Although nonparticipants also considered potential individual benefits such as a ‘medical check-up’ or gaining ‘part of the profit’, they did not especially value or see the individual benefits gained from participating in Lifelines. This lack of benefits for themselves did not motivate them to participate in the biobank. One non-participant (NP31) explained this as follows:

It must clearly matter in my view. For a multi-year research project, one can claim that it is of use. Of course, it is of some use. However, it is difficult to grasp how long it will [continue to] be of use.. .. Of course, if so much data is compared, your data will disappear in a large dataset.. .. Then you get research results from which you must filter out the personal implications. In that sense, for the larger group, it is wonderful that many people are participating. Nevertheless, at this point, I am not interested in participating for this reason [lack of individual benefit].

Whereas some participants and ex-participants expressed the view that personal health experiences featured in their decision to participate in Lifelines, some of the non-participants explicitly stated that they had not had any immediate experiences relating to health issues themselves or within their families. Therefore, they did not perceive the active contribution of personal data as being imperative. As one non-participant (NP23) explained, ‘I don’t know any people with chronic diseases... I need some personal benefit. I can imagine that with Lifelines, you get knowledge about your own health and body. So either that or I am helping someone within my circle’.

Moreover, non-participants mentioned that the burden of investing time for each visit, and for the duration of participation, and the effort expended in relation to the complexity of the performed tasks were also motives for their non-participation. Non-participants found it difficult to acquire an overview of these burdens or costs and of the implications of participating in a population-based biobank. Some of them were hesitant about participating because they were apprehensive about making a long-term commitment to such a project.

Discussion

In this study, we investigated and compared the views of participants and non-participants on contributing to a large-scale, centralized health research data repository with the perceptions of ex-participants who decided to end their participation. We explored the motives, values, and expectations of ex-participants, participants, and nonparticipants of a Dutch

population-based biobank encompassing three generations. Participants and nonparticipants evidenced different perceptions on how they could help others through the contribution of their personal data for biomedical research. Nevertheless, they shared similar intentions about contributing to society and science and helping family members and friends. Participants and non-participants did not agree on realistic risk levels relating to participation. Thus, non-participants and ex-participants expressed more concern relating to participation in large-scale, centralized research data repositories as well as threats of public institutions to constitutional rights. Several studies conducted in Europe and in the United States have confirmed such differences between participants and non-participants in terms of attitudes towards risks and levels of trust, especially in relation to genetic research [19, 23, 26, 32, 45–47].

Our findings showed that intrapersonal differences might exist between participants and non-participants relating to their considerations on trust in society, and in data repositories, as well as their perceptions about society and research. These findings therefore provide some support for significance of the relationship between trust and willingness to participate in research that is well-established [18, 19, 23, 26, 45–47]. Although a lack of trust is associated with high levels of concern and low levels of awareness of current research and data-sharing or data-linking practices [18, 19, 47, 48], our findings indicate that simply increasing the provision of information on data management or biobanking might indeed not be enough [18].

Specifically, our findings show that a lack of engagement in health research practices and inadequate communication of information do not fully explain non-participation in biobanking, as perceptions about society and research appeared to vary between participants and nonparticipants. Recent quantitative studies conducted on individuals unwilling to contribute to biobanks confirmed that their willingness to participate depended on their perceptions and interest or familiarity in biobanking [18, 28]. Moreover, the findings of a pan-European quantitative study and of a systematic review revealed that a preference for broad or narrow informed consent depended on attitudes towards biobanking research, such as benefits, concerns, and information needs [19, 48]. These diverse perceptions of benefits and risks are not confined to potential participants. A recent systematic review confirmed that there were marked variations in the perceptions of members of expert groups and organizations regarding pertinent ethical principles and regulatory norms governing the sharing of international health research data [49].

The formulation of objective criteria for the regulation and the use of personal data is therefore essential in biomedical research and for large-scale, centralized data repositories to counter prejudices based on misunderstandings of risks and benefits of biobanking. In other words, the focus of regulation and use of personal data should be on enhancing trustworthiness of data repositories to provide objective information about the biobanking context and addressing concerns of participation and non-participation [50]. This is the reason why criteria relating to organizations' trustworthiness need to be ingrained within current regulatory and informed consent procedures and implemented more widely, especially in the areas of learning health systems and personalized medicine that entail continuously changing data [50]. For example, a study conducted in Finland revealed that the contexts in which donations were made to a blood bank for patients or to biobanking could trigger different concerns and attitudes [46]. The accounting for context becomes even more key in the case of a large-scale, centralized data repository because of the increasing morally ambiguous and indirect consequences of participation in big data health research for society and for participants [6, 51]. For example, debates about insights in genetic test results and providing feedback to individuals and their families about the genetic risks are ongoing, given differing views on the 'soft' impacts of genetic risks considered as a good practice [52, 53]. An effective regulatory framework should clearly stipulate the responsibilities of data repositories in relation to health research while outlining the implications of participation and non-participation. The findings of several European studies confirm that the trustworthiness of the biobank and of the biobanking context critically influence individuals' decisions to contribute their personal data [19, 32, 46, 48].

A practical implication that follows from a more nuanced understanding of the motives behind participation and non-participation in large-scale, centralized data repositories relates to the need to acknowledge the important influence of intrapersonal factors, such as values, trust in society, and interest. The adoption of a more personalized approach in relation to potential participants could be facilitated through the development of new models for communicating information and recruiting participants for health research [54, 55]. A limited focus on facts and consent within the prevailing generalized information and communication procedures leads to the oversimplification of researcher-participant interactions [8, 52] that does not address the need for more layered and informative interactions in the contemporary context of data collection

and research. Our findings revealed that issues of individual or societal benefits and risks importantly influenced the respective decisions of ex-participants and non-participants to withdraw or not to participate at all. The findings of a recent systematic review of similar studies showed that the public welcomes information about current research and data sharing practices [48]. Therefore, training programmes developed for personnel who recruit individuals to contribute to large-scale, centralized data repositories along with the provision of a comprehensive explanation of current practices and potential future implications of participation to participants and non-participants are important future tasks.

A notable strength of the present study lies in its in depth investigation of motives for (non-)participation. Moreover, by including participants, non-participants, and ex-participants in our study, we have been able to provide a broad overview of the motives for (non-) participation in large-scale, centralized health research data repositories. However, this study also has some limitations. Firstly, the samples, especially of ex-participants, were relatively small. Consequently, some important motives for ending participation may have been missed despite the samples' heterogeneity. For example, it is likely our study lacks the voices of ex-participants and non-participants not consenting due to limited time available or a strong aversion of scientific research. Secondly, we only investigated (non-)participation pertaining to a largescale, centralized data repository focusing on health research. Therefore, our findings and conclusions are bound to the specific context of health research. Their applicability to other settings, such as sustainability research or consumer services, requires further investigation.

Conclusion

In conclusion, we investigated and compared the motives underlying participation, non-participation, and withdrawal of participation in the context of a large-scale, centralized population-based biobank in the Netherlands. Our findings revealed that motives for (non-) participation are complex and that intrapersonal characteristics may be important influencing factors. Acknowledgement of this complexity and the pursuit of personalized approaches relating to recruitment and information communication procedures, along with effective regulatory frameworks, could strengthen collaborative research and healthcare initiatives.

References

1. European Union. Directive 2016/680 of the European Parliament and the Council of the European Union. Off J Eur Union. 2016;L119:89–131. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC>.
2. Mittelstadt BD, Floridi L. The ethics of big data: current and foreseeable issues in biomedical contexts. *Sci Eng Ethics*. 2016;22:303–41. <https://doi.org/10.1007/s11948-015-9652-2>.
3. Murdoch TB, Detsky AS. The inevitable application of big data to health care. *JAMA*. 2013;309:1351–2.
4. Salerno J, Knoppers BM, Lee LM, Hlaing WWM, Goodman KW. Ethics, big data and computing in epidemiology and public health. *Ann Epidemiol*. 2017;27:297–301. <https://doi.org/10.1016/j.annepidem.2017.05.002>.
5. Platt JE, Jacobson PD, Kardias SLRR. Public Trust in Health Information Sharing: a measure of system trust. *Health Serv Res*. 2017;53:824–45. <https://doi.org/10.1111/1475-6773.12654>.
6. O'Doherty KC, Christofides E, Yen J, Bentzen HB, Burke W, Hallowell N, et al. If you build it, they will come: unintended future uses of organised health data collections. *BMC Med Ethics*. 2016;17:54. <https://doi.org/10.1186/s12910-016-0137-x>.
7. Prainsack B. The “we” in the “me”. *Sci Technol Hum Values*. 2017;43:016224391773613. <https://doi.org/10.1177/0162243917736139>.
8. Rivera SM, Brothers KB, Cadigan RJ, Harrell HL, Rothstein MA, Sharp RR, et al. Modernizing research regulations is not enough: It's time to think outside the regulatory box. *Am J Bioeth*. 2017;17:1–3. <https://doi.org/10.1080/15265161.2017.1328899>.
9. Mooney SJ, Westreich DJ, El-Sayed AM. Epidemiology in the era of big data. *Epidemiology*. 2015;26:390–4. <https://doi.org/10.1097/EDE.0000000000000274>.
10. Ioannidis JPA. Informed consent, big data, and the oxymoron of research that is not research. *Am J Bioeth*. 2013;13:40–2. <https://doi.org/10.1080/15265161.2013.768864>.
11. Khoury MJ, Ioannidis JPA. Medicine. Big data meets public health. *Science*. 2014;346:1054–5.
12. Eyal N. Using informed consent to save trust. *J Med Ethics*. 2014;40:437–44. <https://doi.org/10.1136/medethics-2012-100490>.
13. Dickert NW, Eyal N, Goldkind SF, Grady C, Joffe S, Lo B, et al. Reframing consent for clinical research: a function-based approach. *Am J Bioeth*. 2017; 17:3–11. <https://doi.org/10.1080/15265161.2017.1388448>.
14. European Union. Regulation 2016/679 of the European parliament and the

- Council of the European Union. Off J Eur Union. 2016;L119:1–88. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC>.
15. Prainsack B, Buyx A. *Solidarity in biomedicine and beyond*. Cambridge: Cambridge University Press; 2017. <https://doi.org/10.1017/9781139696593>.
 16. Buyx A, Prainsack B. Solidarity can make a difference: addressing transformations in healthcare, demographics and technological replacement. *Bioethics*. 2018;32:537–40. <https://doi.org/10.1111/bioe.12540>.
 17. Prainsack B, Buyx A. Thinking ethical and regulatory frameworks in medicine from the perspective of solidarity on both sides of the Atlantic. *Theor Med Bioeth*. 2016;37:489–501. <https://doi.org/10.1007/s11017-016-9390-8>.
 18. Middleton A, Milne R, Thorogood A, Kleiderman E, Niemiec E, Prainsack B, et al. Attitudes of publics who are unwilling to donate DNA data for research. *Eur J Med Genet*. 2019;62:316–23. <https://doi.org/10.1016/j.ejmg.2018.11.014>.
 19. Gaskell G, Gottweis H, Starkbaum J, Gerber MM, Broerse J, Gottweis U, et al. Publics and biobanks: pan-European diversity and the challenge of responsible innovation. *Eur J Hum Genet*. 2013;21:14–20. <https://doi.org/10.1038/ejhg.2012.104>.
 20. Frizzo-Barker J, Chow-White PA, Charters A, Ha D. Genomic big data and privacy : challenges and opportunities for precision medicine. *Comput Support Coop Work*. 2016;25:115–36. <https://doi.org/10.1007/s10606-016-9248-7>.
 21. Steinsbekk KS, Ursin LØ, Skolbekken JA, Solberg B. We’re not in it for the money-lay people’s moral intuitions on commercial use of “their” biobank. *Med Health Care Philos*. 2013;16:151–62. <https://doi.org/10.1007/s11019-011-9353-9>.
 22. Critchley C, Nicol D, Otłowski M. The impact of commercialisation and genetic data sharing arrangements on public trust and the intention to participate in biobank research. *Public Health Genomics*. 2015;18:160–72. <https://doi.org/10.1159/000375441>.
 23. Nicol D, Critchley C, McWhirter R, Whitton T. Understanding public reactions to commercialization of biobanks and use of biobank resources. *Soc Sci Med*. 2016;162:79–87. <https://doi.org/10.1016/j.socscimed.2016.06.028>.
 24. Ridgeway JL, Han LC, Olson JE, Lackore KA, Koenig BA, Beebe TJ, et al. Potential Bias in the Bank: What Distinguishes Refusers, Nonresponders and Participants in a Clinic-Based Biobank? *Public Health Genomics*. 2013;16: 118–26. <https://doi.org/10.1159/000349924>.
 25. Nissenbaum H. Privacy as contextual integrity. *Washingt Law Rev*. 2004;79: 119–39. <https://digitalcommons.law.uw.edu/wlr/vol79/iss1/10>.
 26. Gaylin DS, Moiduddin A, Mohamoud S, Lundeen K, Kelly JA. Public attitudes about health information technology, and its relationship to health care

- quality, costs, and privacy. *Health Serv Res.* 2011;46:920–38. <https://doi.org/10.1111/j.1475-6773.2010.01233.x>.
27. Kaufman DJ, Murphy-Bollinger J, Scott J, Hudson KL. Public opinion about the importance of privacy in biobank research. *Am J Hum Genet.* 2009;85: 643–54. <https://doi.org/10.1016/j.ajhg.2009.10.002>.
 28. Sanderson SC, Brothers KB, Mercaldo ND, Clayton EW, Antommaria AHM, Aufox SA, et al. Public attitudes toward consent and data sharing in biobank research: a large multi-site experimental survey in the US. *Am J Hum Genet.* 2017;100:414–27. <https://doi.org/10.1016/j.ajhg.2017.01.021>.
 29. Sorani MD, Yue JK, Sharma S, Manley GT, Ferguson AR, Cooper SR, et al. Genetic data sharing and privacy. *Neuroinformatics.* 2014;13:1–6. <https://doi.org/10.1007/s12021-014-9248-z>.
 30. Verlinden M, Nys H, Ectors N, Huys I. Qualitative study on custodianship of human biological material and data stored in biobanks. *BMC Med Ethics.* 2016;17:15. <https://doi.org/10.1186/s12910-016-0098-0>.
 31. Joly Y, Dalpé G, So D, Birko S. Fair shares and sharing fairly: A survey of public views on open science, informed consent and participatory research in biobanking. *PLoS One.* 2015;10:1–21. <https://doi.org/10.1371/journal.pone.0129893>.
 32. Middleton A, Bragin E, Morley KI, Parker M. Online questionnaire development: Using film to engage participants and then gather attitudes towards the sharing of genomic data. *Soc Sci Res.* 2014;44:211–23. <https://doi.org/10.1016/j.ssresearch.2013.12.004>.
 33. Lemke AA, Wolf WA, Hebert-Beirne J, Smith ME. Public and biobank participant attitudes toward genetic research participation and data sharing. *Public Health Genomics.* 2010;13:368–77. <https://doi.org/10.1159/000276767>.
 34. Murad AM, Myers MF, Thompson SD, Fisher R, Antommaria AHM. A qualitative study of adolescents' understanding of biobanks and their attitudes toward participation, re-contact, and data sharing. *Am J Med Genet Part A.* 2017;173:930–7. <https://doi.org/10.1002/ajmg.a.38114>.
 35. Broekstra R, Aris-Meijer J, Maeckelberge E, Otten S, Stolk R. Trust in Centralised Large-Scale Data Repository: a qualitative analysis. *J Empir Res Hum Res Ethics.* 2019. <https://doi.org/10.1177/1556264619888365>.
 36. Scholtens S, Smidt N, Swertz MA, Bakker SJLL, Dotinga A, Vonk JM, et al.

- Cohort profile: LifeLines, a three-generation cohort study and biobank. *Int J Epidemiol.* 2015;44:1172–80. <https://doi.org/10.1093/ije/dyu229>.
37. Ziebland S, McPherson A. Making sense of qualitative data analysis: an introduction with illustrations from DIPEX (personal experiences of health and illness). *Med Educ.* 2006;40:405–14. <https://doi.org/10.1111/j.1365-2929.2006.02467.x>.
 38. Kvale S. *Doing interviews*. Los Angeles: Sage Publications; 2007. <https://doi.org/10.4135/9781849208963>.
 39. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant.* 2018;52:1893–907. <https://doi.org/10.1007/s11135-017-0574-8>.
 40. Coyne IT. Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries? *J Adv Nurs.* 1997;26:623–30. <https://doi.org/10.1046/j.1365-2648.1997.t01-25-00999.x>.
 41. Baker SE, Edwards R. How many qualitative interviews is enough? *Natl Cent Res Methods Rev Pap.* 2012;1–42. <https://doi.org/10.1177/1525822X05279903>.
 42. Mason M. Sample size and saturation in PhD studies using qualitative interviews. *Forum Qual Sozialforsch / Forum Qual Soc Res.* 2010;11:19. <https://doi.org/10.17169/fqs-11.3.1428>.
 43. Layder D. *Sociological practice: linking theory and social research*; 1998. [http://books.google.nl/books?hl=nl&lr=&id=hTgzIESnrewC&oi=fnd&pg=PP2&dq=Layder,+D.+\(1998\).+Sociological+practice+:+Linking+theory+and+social+research.+London:+Sage.&ots=po2RJFBImE&sig=RmjXMGMQbfcFyzaVUoGsoN99QdE](http://books.google.nl/books?hl=nl&lr=&id=hTgzIESnrewC&oi=fnd&pg=PP2&dq=Layder,+D.+(1998).+Sociological+practice+:+Linking+theory+and+social+research.+London:+Sage.&ots=po2RJFBImE&sig=RmjXMGMQbfcFyzaVUoGsoN99QdE). Accessed 28 Nov 2016.
 44. Friese S. *ATLAS. Ti 7 user guide and reference*; 2013. p. 1–469.
 45. Kettis-Lindblad Å, Ring L, Viberth E, Hansson MG. Genetic research and donation of tissue samples to biobanks. What do potential sample donors in the Swedish general public think? *Eur J Public Health.* 2006;16:433–40. <https://doi.org/10.1093/eurpub/cki198>.
 46. Raivola V, Snell K, Helén I, Partanen J. Attitudes of blood donors to their sample and data donation for biobanking. *Eur J Hum Genet.* 2019;27:1659–67. <https://doi.org/10.1038/s41431-019-0434-1>.

47. Milne R, Morley KI, Howard H, Niemiec E, Nicol D, Critchley C, et al. Trust in genomic data sharing among members of the general public in the UK, USA, Canada and Australia. *Hum Genet.* 2019;138:1237–46. <https://doi.org/10.1007/s00439-019-02062-0>.
48. Aitken M, De St Jorre J, Pagliari C, Jepson R, Cunningham-Burley S. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. *BMC Med Ethics.* 2016;17:1–24. <https://doi.org/10.1186/s12910-016-0153-x>.
49. Kalkman S, Mostert M, Gerlinger C, Van Delden JJM, Van Thiel GJMW. Responsible data sharing in international health research: a systematic review of principles and norms. *BMC Med Ethics.* 2019;20:21. <https://doi.org/10.1186/s12910-019-0359-9>.
50. O'Neill O. Linking trust to trustworthiness. *Int J Philos Stud.* 2018;26:293–300. <https://doi.org/10.1080/09672559.2018.1454637>.
51. Verheij RA, Curcin V, Delaney BC, McGilchrist MM. Possible sources of bias in primary care electronic health record data use and reuse. *J Med Internet Res.* 2018;20:e185. <https://doi.org/10.2196/jmir.9134>.
52. Kraft SA, Cho MK, Gillespie K, Halley M, Varsava N, Ormond KE, et al. Beyond consent: building trusting relationships with diverse populations in precision medicine research. *Am J Bioeth.* 2018;18:3–20. <https://doi.org/10.1080/15265161.2018.1431322>.
53. Jackson C, Gardy JL, Shadiloo HC, Silva DS. Trust and the ethical challenges in the use of whole genome sequencing for tuberculosis surveillance: a qualitative study of stakeholder perspectives. *BMC Med Ethics.* 2019;20:1–14. <https://doi.org/10.1186/s12910-019-0380-z>.
54. Kaye J, Whitley EA, Lund D, Morrison M, Teare H, Melham K. Dynamic consent: a patient interface for twenty-first century research networks. *Eur J Hum Genet.* 2014;23:141–6. <https://doi.org/10.1038/ejhg.2014.71>.
55. Ploug T, Holm S. Meta consent: a flexible and autonomous way of obtaining informed consent for secondary research. *BMJ.* 2015;350:h2146. <https://doi.org/10.1136/bmj.h2146>

Chapter 3

Trust in Large-Scale, Centralized Data Repository: A Qualitative Analysis

Abstract

Exponential increases in digital data and calls for participation in human research raise questions about when and why individuals voluntarily provide personal data. We conducted 36 in-depth interviews with ex-participants, participants, and nonparticipants in a biobank to identify key factors influencing trust in centralized large-scale data repository for human research. Our findings indicated that trust depends strongly on whether such data repository benefits the public, the interests of data collectors, the characteristics of the collected data, and application of informed consent for retaining control over personal data. Concerns about the aims and range of data repository appeared to influence withdrawal of participation. Our findings underscore ethical and practical issues relating to data collection and consent procedures in human research.

Keywords: biobanks; participation; withdrawal; data linkage; commercialization trust, big data, biorepositories/biobanks, cohort study, decision making, qualitative methods, the Netherlands, justice/ participant selection/inclusion/ recruitment

This chapter is previously published as: Broekstra R, Aris-Meijer J, Maeckelbergh E, Stolk R, Otten S. Trust in Centralized Large-Scale Data Repository: A Qualitative Analysis. Journal of Empirical Research on Human Research Ethics. 2020;15(4):365-378. doi:10.1177/1556264619888365

Introduction

In recent years, there has been an exponential increase in digital data and calls for participation in human research through the provision of personal information for large-scale, centralized data repositories, especially in the medical sciences. This phenomenon warrants an investigation of when and why individuals voluntarily provide data or withdraw their participation (Broekstra et al., 2017). The overall increase in digital data repository has raised concerns among experts as well as participants in centralized largescale data repository about autonomy, privacy, and responsibility (Mittelstadt et al., 2016; Mittelstadt & Floridi, 2016; Sorani et al., 2015; Stricker, 2017). Innovative methods applied in large-scale, centralized data repository and the linking of personal data, such as blood samples, to other types of health and nonhealth information permit an even wider range of data use than do traditional research methods (Prainsack & Buyx, 2013; Steinsbekk, Ursin, et al., 2013).

Such data usage, however, deviates from the ambit of current standards for safeguarding participants and nonparticipants against violations of integrity, such as standard informed consent and anonymization procedures, and methods of protecting data (Barocas & Nissenbaum, 2014; boyd & Crawford, 2012; Keymolen, 2016; Mittelstadt & Floridi, 2016; Robinson et al., 2013). Moreover, it raises questions regarding the confidentiality of personal data, possible stigmatization, and uncertainty about who has access to the data (boyd & Crawford, 2012; Gibson et al., 2017; Graeff & Harmon, 2002; Mittelstadt & Floridi, 2016; Rivera et al., 2017). Health and medical care will be prioritized in relation to science and innovation over the next 15 years (Eurobarometer, 2014). Therefore, a study of the different perspectives of stakeholders in human research entailing large-scale, centralized data repositories on data sharing and trust is pertinent. Specifically, we compared the perspectives of ex-participants, participants, and nonparticipants in biobanks for facilitating human research in the medical sciences.

With the establishment of biobanks, concern has shifted from questions of privacy to those of trust (Kelley et al., 2015; Mittelstadt & Floridi, 2016). Biobanks store large quantities of biological specimens as well as data extracted from questionnaires and measurements. They are aimed at facilitating studies on patients with specific diseases or prospective studies on the onset and progress of chronic diseases (Krokstad et al., 2013; Scholtens et al., 2015; Sudlow et al., 2015; UK Biobank Coordinating Centre, 2007; van Staa et al., 2016). Studies of biobanks, entailing extensive data repository over a long duration, differ from small-scale or cohort studies. Centralized information storage in biobanks in-

evitably poses a threat to participants' privacy. The decision to participate in a biobank entails accepting unknown short-term and long-term privacy risks concerning the confidentiality of personal data, the use of such data for unintended research purposes, or possible stigmatization (boyd & Crawford, 2012; Gibson et al., 2017; Graeff & Harmon, 2002; Mittelstadt & Floridi, 2016; Nobile et al., 2016; Rivera et al., 2017). This situation results in a social dilemma in which the collective interest of promoting the public good and the individual interest of keeping data or information private are opposed. Trust may reduce the complexity of this dilemma through general assumptions whereby an unfamiliar world or future is simplified, thus enabling decision making on risk taking (Luhmann, 1979). Studies have shown that trust in biobanks, health care providers, and other public institutions is a key factor in determining individuals' willingness to participate in epidemiological research and to provide personal data to biobanks (Critchley et al., 2012; Gaskell et al., 2013; Hansson, 2005; Kaufman et al., 2009; KettisLindblad et al., 2006; Lemke et al., 2010; Nobile et al., 2016; Rahm et al., 2013).

Currently, biobanks and other facilities used in human research attempt to enhance or change the complexity, volume, and nature of their data repositories to optimize them (Mittelstadt & Floridi, 2016). Researchers and data collectors seek to link individuals' data to other types of personal data available within analyses of large data repositories to illuminate complex patterns and build predictive models (Barocas & Nissenbaum, 2014; Nuffield Council on Bioethics, 2015). These "big health data" can emerge out of detailed data revealing individuals' characteristics, behaviors, and preferences, for example, through combined analyses of questionnaires and genomic, register, sensory, satellite, and social media data. Accordingly, voluntary participation in research in a context of openly shared information requires a greater degree of trust in the data repositories and their use in human research, compared with a context in which enrichment or sharing of data does not occur (Gibson et al., 2017; van Staa et al., 2016).

Trust, however, is a concept that is complex and difficult to grasp (Parks et al., 2015). Although trust seems to be a requirement for participation in research, its function within decision making regarding participation or withdrawal remains unclear. Moreover, the distinction between trust and distrust is still a matter of debate, with some scholars viewing trust and distrust as two ends of a single continuum, whereas others view them as distinct (Keymolen, 2016; Luhmann, 1979; Saunders et al., 2014). Trust can be defined as an individual's willingness to be vulnerable without evaluating the behavior and intention of another person that automatically leads to a decision on whether or not to accept the risk (Ferrin et al., 2006; Mayer et al., 1995). Other scholars regard trust as a cognitive

construct relating to the outcome of an individual's reasoning process regarding their own vulnerability when evaluating others' behaviors and intentions (Bhattacharya et al., 1998; Lewicki et al., 1998). The contradictory findings of studies on trust and biobanking lend support to both views. According to the findings of a study conducted in the United States in 2009 ($n = 4,659$ respondents), regardless of trust, privacy remains an important issue for potential participants in biobank studies, indicating the prevalence of automatic judgments (Kaufman et al., 2009). However, the findings of another study indicating that privacy concerns were less salient in decisions to participate in a biobank study when high levels of generalized trust existed support a conception of trust as a cognitive rational construct (Gaskell et al., 2013).

It is therefore important to understand why and under what conditions individuals are willing to trust a biobank with their personal data and to identify specific conditions that contribute to a reduction of this trust (Nobile et al., 2016; Petersen et al., 2014). In particular, ex-participants' reasons for withdrawal can yield valuable insights into the process of trust building. A comparative study encompassing ex-participants, participants, and nonparticipants in a biobank that has introduced new methods for optimizing data repositories in an open context of information sharing would advance knowledge regarding the pathways that lead to the acquisition or loss of trust. Several studies have focused individually on ex-participants, participants, and nonparticipants (Critchley et al., 2016; Gaskell et al., 2013; Nobile et al., 2013; Steinsbekk, Ursin, et al., 2013), although empirical investigations of ex-participants are somewhat limited (Ridgeway et al., 2013). However, few studies have examined these groups simultaneously. Therefore, we explored the conditions under which individuals voluntarily entrust a biobank with their personal data and the reasons why trust sometimes decreases over time.

Method

Data Collection

We conducted semi-structured individual interviews, as they enable an in-depth exploration of respondents' perceptions, events, and experiences relating to centralized largescale data repository and research (non)participation. Moreover, their opinions and statements can be triangulated and clarified (Kvale, 2007, p. 11). We developed an interview topic guide and applied a narrative approach that left room for discussion of unanticipated themes. This narrative approach was partly derived from the method used in the DIPEX project (Ziebland & McPherson, 2006) that focused on personal experiences of health and illness.

The interview topic guide, shown in Table 1, was developed on the basis of six current themes identified in the scientific literature on (non)participation in biobanking, public goods, trust, and data sharing (Balliet & Van Lange, 2013; D’Abramo, 2015; Gaskell et al., 2013; Kaufman et al., 2009; Mittelstadt et al., 2016; Mittelstadt & Floridi, 2016; Nobile et al., 2016, 2013; Parks et al., 2013; Sorani et al., 2015; Steinsbekk, Myskja, & Solberg, 2013). The following themes were identified: (a) becoming a (non)participant, (b) objective aspects of participation (e.g., general feelings, tasks, and feedback), (c) subjective aspects of participation (e.g., expectations of accomplishment and feelings of identification), (d) understandings of and attitudes toward centralized data repository and linkage, (e) perceived benefits and threats relating to centralized data repository and linkage, and (f) decisions on whether or not to provide personal data. The interview topics were subsequently derived from these themes.

Table 1. Interview topics and illustrative questions

Topic	Example questions
Association with biobank	How did you get involved?
Decision of (non-)participation	What were your considerations?
Reality	Was your experience as expected or not?
Future	Will you participate in Lifelines or in another centralised large-scale research project in the future?
Overall perception	How is participation in Lifelines perceived in general?
Tasks	How are the tasks related to participation in Lifelines perceived?
Feedback	What do you think of the feedback on your data you (would) receive?
Outcome hoped for	What motivates individuals to participate?
Benefits	What are the benefits and opportunities relating to participation?
Disadvantages	What are the risks and disadvantages relating to participation?

Definition	What are centralised large-scale data repositories?
Attitudes towards large-scale centralised data repositories	Are the public's attitudes towards collection, linking and use of this data generally positive or negative?
Benefits of large-scale centralised data repositories	What are the benefits of collecting, linking and using data?
Risks posed by large-scale centralised data repository	What risks are posed by the collection, linking and use of data?
Personal balance of benefits and threats	From your personal perspective, how are these considerations balanced?
Societal balance of benefits and threats	From a societal perspective, how are these considerations balanced?

All of the interviews were conducted by one member of the research team with training and experience in interviewing techniques (R.B.). Current participants in Lifelines, a Dutch population-based biobank, were interviewed following their regular visits to the Lifelines research site. Most of the interviews were held at this site or in a room allocated for the interviews at the University Medical Center Groningen. All four ex-participants from Lifelines and eight of the fifteen nonparticipants were interviewed at their homes at their request. Each interview covered all of the topics in the guide and lasted between 30 and 65 min, depending on when saturation was reached in interviews. The interviews were audio-recorded with the consent of the interviewees. Only one interviewee denied permission for audio recording the interview, but allowed notes to be taken. All of the recordings were transcribed by an independent professional organization. The complete study design protocol received ethical approval from the Medical Ethics Review Board.

Recruitment and Sampling

Maximum variation sampling was applied to recruit interviewees (Coyne, 1997). Consequently, a heterogeneous sample was obtained (Baker & Edwards, 2012; Mason, 2010). Between August and September 2016, a total of 36 individuals were interviewed: 17 had participated in Lifelines, 4 were ex-participants, and 15 did not participate in Lifelines. In this article, we refer to these distinct groups of interviewees as “participants,” “ex-participants,” and “nonparticipants.” Table 2 shows the characteristics of each of the interviewees. The average age of the interviewees was 45 years (a range of 20–68), and 17 interviewees (47%) were male.

Table 2. Respondents' Characteristics

Respondent	Sex	Age	Occupational Field	Participation status
1	Female	67	Human resources	Participant
2	Male	42	Transportation	Non-participant
3	Female	60	Education	Non-participant
4	Male	30	Legal and social work	Participant
5	Male	52	Photography and architecture	Ex-participant
6	Male	60	Municipality	Non-participant
7	Female	37	Home maintenance	Ex-participant
8	Male	26	Warehouse maintenance	Participant
9	Female	59	Retail	Participant
10	Male	33	Entrepreneur	Participant
11	Male	20	Student	Participant
12	Female	53	Law	Participant
13	Male	30	Student	Non-participant
14	Male	66	Mechanical engineering	Participant
15	Male	55	Home maintenance	Non-participant
16	Female	29	Publisher	Non-participant
17	Female	50	Entrepreneurship	Non-participant
18	Female	68	Retired	Participant
19	Male	54	Management	Participant
20	Female	27	Energy policy	Non-participant
21	Female	39	Healthcare	Participant
22	Female	50	Healthcare	Participant
23	Female	27	Human resources	Non-participant
24	Female	46	Education	Participant
25	Male	45	Municipality	Participant
26	Female	65	Farming	Non-participant
27	Female	43	Librarian	Ex-participant
28	Female	51	Social work	Ex-participant
29	Female	37	Management	Participant
30	Male	33	Engineering	Non-participant
31	Male	60	Human resources	Non-participant
32	Female	33	Home maintenance	Non-participant
33	Male	45	Management	Participant
34	Female	63	Retired	Participant
35	Male	34	Business development	Non-participant
36	Male	49	Entrepreneurship	Non-participant

We recruited interviewees partly from the Lifelines biobank. Lifelines is a multidisciplinary population-based prospective cohort study focusing on the health and health-related behaviors of persons who have been living in the Northern Netherlands since 2006. The study has a unique three-generational design and employs a broad range of investigative procedures for assessing the biomedical, sociodemographic, behavioral, physical, and psychological factors that collectively contribute to the health of the general population and to incidences of disease, with a special focus on multimorbidity and complex genetics (Scholtens et al., 2015). The recruitment of new participants for the Lifelines project stopped in 2013. To ensure their privacy, participants and ex-participants were recruited by the Lifelines organization. Ex-participants who had ended their Lifelines's participation in the last 2 months prior to the commencement of our study in August 2016 were invited by phone to participate in our study. The invitation was communicated by Lifelines in the last stage of the withdrawal procedure of Lifelines, but the invitation protocol was written by the researchers of this study (R.B. and J.A.). Moreover, nonparticipating partners of the Lifelines participants were invited by the organization. Ex-participants were the least willing to participate in our study. Although we invited ninety such individuals, only four consented to participate (4.44%). This response rate was significantly lower than those of participants and nonparticipants (22.2% and 24.9%, respectively). The reasons given by invitees for not participating in our study were "no interest," "not available," and "not willing to invest more time in Lifelines or its studies." Although our sample of ex-participants is too small to allow for generalizations, we nevertheless obtained the in-depth information that we sought. All nonparticipants were members of the public whom we recruited through personal face-to-face invitations issued at the entrance of a central public library and through personal or general written invitations that followed the same invitation protocol issued via various social networks as well as on online and offline platforms. Our recruitment and sampling was in accordance with the approval of the Medical Ethics Review Board.

Data Analysis

We analysed each individual phrase of all 36 interview transcripts in the context of the overall interview content, and, where appropriate, we generated or assigned content-related codes. Codes could be applied multiple times in a single transcript, and phrases could consist of multiple codes. Codes with related content were grouped together, and these groups were subsequently categorized into themes. We developed an initial coding protocol after conducting a close reading of three transcripts. This coding protocol was evaluated by two

or three researchers working on each transcript through an iterative process of cross-checking the individual transcript analyses. After consensus had been reached about the content of codes, groups, and themes, the resulting coding protocol was finalized and used for the remaining transcripts. There was overall agreement on the coding. When different codes were used, consensus was reached through discussion. Five researchers were involved in coding the transcripts. Six random transcripts were coded by at least three researchers. Subsequently, one researcher coded all of the transcripts, whereas the other researchers coded a limited number of transcripts. In line with the protocol, the results of the interviews were categorized into four themes for ex-participants (EP), participants (P), and nonparticipants (NP). One of these themes, "Attitude toward centralized data repository and linkage in the context of biobanking," was influenced by our topic guide, given the focus of discussions during the interviews prior to the analysis. The prominence of the other themes did not emerge directly from our topic guide but was revealed by the data. The validity of this process of identifying themes is confirmed by the adaptive theory on qualitative research, which allows for such theoretical influence within research (Layder, 1998). Transcripts were primarily analysed using a qualitative data analysis software package, Atlas TI version 7, to enable the retracing and evaluation of quotes along with their codes, groups, and themes (Frieze, 2013).

Results

We identified multiple conditions that relate to experiencing the trust required to provide personal data in the context of biobanking and large-scale, centralized data repository. However, interviewees found it difficult to explain the reasons for this trust. For example, one interviewee (NP17) provided the following explanation: "I can't express it in words; it's a feeling." Our analysis of the interview transcripts revealed four major themes that featured in the formation of trust and in decisions made by individuals about sharing their data. The first theme that emerged from the data of unique quotes was data repository aimed at benefiting the public. A second theme was organizational aspects of the biobank as a data collector, focusing on how the biobank is organized to provide a context in which data are used safely. The third theme was characteristics of the collected data, which concerned the type of data requested and its relevance for research. The last theme was control over personal data maintained through informed consent. This theme related to the influence of interviewees' perceptions of informed consent in line with their intention of maintaining control over the personal data that they provided.

Data Collection Aimed at Benefiting the Public

The interviewees characterized the context of medical scientific research as a quest aimed at improving public health. Although ex-participants as well as several nonparticipants expressed a lack of trust in public goods, and especially in medical scientific research, participants and the majority of nonparticipants perceived the context of medical scientific research to be generally trustworthy and affirmed their trust in Lifelines as a public good. For example, Lifelines was widely perceived as a charitable organization, as observed by one nonparticipant (NP13): “If I understand the question correctly, for me it is actually almost [like] a charity, in which I would participate. [But] instead of money, I give my time.” Participants placed considerable value on the objective of benefiting the public. Several interviewees, mostly participants, stated that they would be unwilling to provide their personal data if the research was primarily aimed at providing financial gains for just a small group of people, such as the managers or shareholders of Lifelines. One participant (P4) stated, “Well, I think it’s different. If it is for a [for-] profit organization then I think . . . look, that’s not why I participated in it.” Another participant (P14) provided the following explanation:

I see someone who wants to sell a washing machine differently from someone who wants to collect data about a certain disease in a certain family. Because the first one tries to sell a washing machine through data [received] from my laptop. That data is passed on just for personal gain. He wouldn’t otherwise [ask for it]. I don’t have that feeling with Lifelines [as data repository]. Lifelines aims to make someone better, but then your offspring instead of yourself.

By contrast, most nonparticipants and a few participants were less concerned about this objective. They expressed their willingness to share their data in return for personal financial gain if they were interested in the research topic. This was expressed by one of these participants (NP2) as follows:

Yes, if they have a commercial interest, I want them to share their profit as well. Then I’m going to look at it in a different way. If you were to make a profit, I would like you to say so.

Organizational Aspects of the Biobank as a Data Collector

All four ex-participants stated that they had withdrawn their participation because they were disappointed to some extent in the organizational aspects of the biobank, which contributed to a reduced level of trust. For three of four of the ex-participants, the extensive scale of data repository was a critical factor influencing their distrust because they worried about the use of their personal data in Lifelines and the preservation of their anonymity within such a large-scale project. They perceived large-scale, centralized data repository to be valuable for researchers, but this also entailed an increased risk. Two respondents explained,

I believed that it would only be used within the hospital or at the University Medical Center. But it goes further in my opinion. . .

. See, when you develop medicine or something like that, you need information about large groups of people. I've heard that even secret services and the police are interested in DNA material. There is no other large biobank in existence apart from this one with all of its information. From a commercial perspective, this is of course very interesting. (EP5)

Well I believe your data [at Lifelines] are no longer anonymous. I just don't believe that [data are anonymous] anymore. That doesn't concern me a lot. Only you don't know which way things are going. You never know where your [medical] data is going to end up and how it will be used or valued. (EP27)

All participants and most nonparticipants believed, however, that the larger renowned institutes were more competent than smaller institutes because they perceived them to be “better equipped,” “more solidly financed,” and “more experienced in handling their participants’ information.” They did not worry about privacy threats within Lifelines because of the substantial number of participants who would be affected if data were to be leaked or misused. As one participant (P10) observed, “The bigger it gets, the more anonymous [it becomes].” They assumed that data were secure. As one nonparticipant (NP16) explained,

No, well I just assume that the information will be anonymous and wouldn't become public. That is something you may expect not to happen. You need to take care of security breaches or data leakages when designing such large-scale research.

Moreover, the ex-participants had expected more effective communication regarding privacy, aims, and feedback on their results. Their perception of the lack of proactive communication of the organization and its representatives was the final factor prompting their distrust. One ex-participant (EP7) described,

You have to hand in blood and urine, but you never get an answer [feedback on results]. Yes, if it's really bad, they will answer, but by then it is too late. Otherwise they won't give me an answer, for example "That's good and that's for the lungs and the heart." They don't answer that. You don't get an answer on whether the heart is still good or not.

All of the participants and nonparticipants indicated that their trust was also based on the anticipated quality and motives of the employees or other representatives of the organization. They mentioned that representatives' competencies, measured, for example, by the expected university degree qualification, influenced levels of trust. Moreover, "professionalism" signaled trustworthiness. Such perceptions of "professionalism" were based on the formal and consistent behavior of the biobank representatives (e.g., being friendly and calm). This was explained by one respondent (P8) as follows: "It is true that they always take a look at you; they are interested, so to speak. I think that it is very important to show an interest in the people who are visiting."

None of the ex-participants, participants, or nonparticipants had a comprehensive understanding of how security measures and purpose limitations for data repositories are applied in medical scientific research in general, and specifically by Lifelines. This was a factor influencing the development of distrust or trust among interviewees. In light of their negative experiences that reinforced their personal beliefs, the ex-participants (EP7 and EP8) associated Lifelines with poor medical research and care and with (pharmaceutical) business (EP5 and EP27). As one of them (EP7) explained, this perception was based on information acquired indirectly:

I heard rumours in the newspapers and so forth that some general practitioners had stated that one should not participate in it. Moreover, I heard via an acquaintance who works in IT and was involved partly [in Lifelines] that it works a bit like that [business].

Neither of the above assumptions is backed by objective evidence. Three of the ex-participants (EP5, EP27, and EP28) and some of the nonparticipants were convinced that Lifelines was unable to guarantee the anonymity of the data, which could result in potential misuse of data by the government or by health insurance companies. As one nonparticipant (NP23) observed,

I wonder what a government does with this kind of information. They are not doctors or hospitals, you know. So what do you want to do with them? I can also imagine that they would like to share this with health insurers or people like that. Then I think, sorry, but if that means that in 10 years' time, for example, I will need enormously expensive medicines, then my health insurer doesn't need to know. So it depends a bit on who and for what purposes.

Lifelines provided participants with information about security measures and strictly regulated methods that were deployed at the onset and during their participation via information leaflets, informed consent forms, and newsletters. However, none of the participants could explain in detail how their donated material was being used, stored, and shared. One participant (P10) stated,

What you see and the information you get, that will be sufficient. Yes, it sounds very naive, but you can't really do it. . . . Going there on one visit, you're not going to ask: "Show me what you're doing." No, you don't do that. You just assume that this is good.

As one participant (P4) clarified, the biobank was perceived as trustworthy because of its link with the University of Groningen and with the University Medical Center Groningen:

I believe that the University Medical Center is the mother of the organisation. That is one of the most important organisations in the medical sciences, here in the North. That scientific element of the research and the organisation is at the top of my list of factors that play a role. Well, if it wasn't scientific, not linked to the university, then I would have to think about it again perhaps.

Characteristics of Collected Data

Interviewees mentioned that the sensitivity of the data that they provided, and hence the level of trust required, was influenced by the range of collected data. Interviewees suggested that the aim of the research should match the data requested. Two ex-participants stated that the broad range of data collected by Lifelines was an important concern for them influencing their withdrawal as participants. In their view, if, in addition to medical details, the data encompassed supplementary information on aspects of their personal lives, such as their financial information, then such data would be more sensitive and misuse of the data could have a significant impact on their lives. According to one ex-participant, it could be explained as follows (EP27): “before that [negative experiences during visits], I had seen a documentary about privacy. So that may have been one of the factors because you don’t know what the consequences of [sharing] your information are.” All of the ex-participants and most of the nonparticipants had concerns about the possible abuse of their data in the future and were wary about providing data to the biobank, especially via data linkages. Some of them felt that a future trend of organizations controlling people’s lives was likely. For example, one nonparticipant (NP15) stated, “But if you have a lot of data, then you can manipulate people. In the end, you can use every piece of information about someone to control him or her.” The distrust of these interviewees mainly stemmed from centralized largescale data repository outside the scope of scientific research, such as data extraction from social media or web browsers, and especially search engines. However, they also mentioned other elements of distrust relating to such data repository in the context of medical scientific research. Thus, a participant (P1) made the following observation: “Well a very good example is that this knowledge will enable you to potentially exclude people from insurance. That is something that frightens me.”

These concerns were also shared by other interviewees to a lesser extent. One nonparticipant (NP16) described the following concern: “If they acquire data on other aspects of our lives, then they truly become ‘Big Brother’ they can really dig into your life.” More specifically, these interviewees were concerned about the possibility of the Lifelines database being linked to data held by institutes that did not focus on health care, for example, Statistics Netherlands or the Dutch Tax authority. This concern is expressed in the following quote:

For example, my tax information does not seem to me to be relevant for Lifelines. That is my intuitive feeling. If someone could explain to me why this information is relevant, that could change my opinion. However, my primary response is that it is not necessary to provide such information. (NP30)

The interviewees who had concerns about large-scale, centralized data repository expressed doubts as to whether their personal data would be used in a way that they perceived as nonthreatening in the future, given Lifelines's linkages with such institutions.

Yes, they can dig into all your facets of life. Of course, medical information is already quite personal. The most personal thing you can know about someone. But if they also want to see your pay check and so on, it might well be a matter of concern. (NP16)

Several nonparticipants and some participants, however, perceived opportunities in addition to risks relating to participation in research and the linking of personal data. Their interest in and support of medical scientific research were decisive factors influencing their consent to participate in this large-scale, centralized data repository. They stated that they would supply a broad range of requested information providing that certain conditions were met, such as adherence to their informed consent and its importance in relation to a relevant research context. One participant (P17) stated that “if they could make it clear to me why that is important, I would consider it. I wouldn’t just say beforehand ‘oh yes, just take it.’” Another participant (P9), who had reservations, observed, “Well, if it is about your well-being and your health, then I think keep [the data request] to that theme.” These participants mentioned having a gut feeling that something could be wrong with data sharing in general or that their privacy could be threatened after discussing their data contributed to the biobank. These apprehensions mainly stemmed from their conception of a future entailing an increasing risk of data being misused by organizations to influence behavior. One participant (P1) provided the following explanation:

Many people do these things and what they will do with that huge mountain [of data]. I fear no one knows anymore what can be done or what cannot be done with it. There are no inputs into this matter, so to speak. So abuse is also obvious or conceivable.

Despite these apprehensions, however, the interviewees felt that the provision and linkage of their data to the biobank was inevitable, given the high degree of relevance of their medical information. A participant (P4) stated,

It is a bit scary. Nowadays they want to know everything about you, and even the government knows where you are, where you drive, where you check in with a public transport card. But yes, that is also the time we are living in. So I think we should go along with that.

None of the other participants had any reservations about the increasing provision and linking of their personal data to the biobank via other organizations because they had nothing to hide within the specific context of human research in medical sciences. The objective of health care justified their full contribution to the biobank and to medical scientific research. Two participants explicitly stated that if Lifelines considered financial information necessary for research purposes, then they would provide it without stipulating any conditions:

How far can they go with that [data collection]? Extremely far for me. if that requires me to reveal sensitive things, like what diseases I have had, how my sex life is or how much I drink, no problem. Even if they put it on the Internet, if I [help to] make someone better with it [the information], then they should do that. (P10)

Another participant (P11) provided the following explanation:

You give away data, and for the rest, you are not 100% sure if anyone will ever leak something . . . or that it will be used against you ever But I don't attach a great deal of importance to that if I compare that with what kind of benefits there can be, so to speak.

Control Over Personal Data Maintained Through Informed Consent

The interviewees' subjective perceptions of informed consent given to the biobank played an important role in their decisions to withdraw their participation or contribute data. All of the ex-participants indicated that their reason for withdrawing was to take back control over their data because they felt that they had been misinformed or given insufficient information. One of them (EP5) stated,

Well, I couldn't find anything on the Internet, for example, from the Lifelines organization on what exactly they do with all the information. That was nowhere to be found. Then I thought, what's secret? . . . I believe that journalists have been looking for that [information] too. I couldn't find anything about it. I think they should disclose that much more. What exactly do they do with it? What exactly is their goal?

Thus, they lost trust not only in the biobank but also in data sharing in general or in health care. Their trust diminished after they were misinformed about data management or the nonprofit policy, or when they developed a distrust of data sharing in general. It also diminished after they gained access to new in-

formation or exposure to media debates about research and data security. Although some interviewees from the group of participants had given their consent, they were still somewhat hesitant to provide their personal data because of a perception of loss of control: “You don’t have any control. My problem concerns what they do with it. You have to have that confidence in them, which I find a bit difficult.” (P3)

One third of the interviewees, mainly participants, mentioned that providing personal data to the biobank would not be an issue because informed consent had been or would be given. They felt that they had adequate control over their donated personal data and could “always” retract their data if necessary:

If I must think about everything, that’s simply not me. “Oh, what is he going to do with my data?” I don’t care. I just don’t say things that I don’t want to tell them. (P22)

Nonparticipants appeared to be more cautious and valued the personal data they would provide as sensitive. For example, one nonparticipant (NP3) explained,

Look, if I am part of a group, and I have to make a choice. . . . I have no control. What concerns me is what they do with it. You must give them that trust, but I find that a bit difficult. In my opinion, it is the same with politics. Municipal politics, I find that easier. Otherwise it’s [politics] too broad.

Although views on the value of the contributed personal data differed, the procedure of informed consent was widely valued as an important symbolic act. Although four interviewees (two participants and two nonparticipants) acknowledged that they themselves could not verify how data were managed at Lifelines, they still preferred to retain some form of control over what happened with and to their data.

Discussion

Our aim in this study was to understand why and how people are willing to entrust their data to a facility engaged in large-scale, centralized data repository for human research and what conditions would contribute to a reduction in their trust. Accordingly, we focused on possible differences in trust levels and in perceptions relating to data sharing among ex-participants, participants, and nonparticipants in a Dutch population-based biobank. Our results showed that evaluations of the biobank were related to its aim of being a public good and its grounding in medical science. Questions about its aims and nature as a public good and concerns about privacy seemed to be key factors influencing the withdrawal of participation and the building of trust. We found that the following four factors influenced trust building relating to large-scale, centralized data repository for human research: data repository aimed at benefiting the public, the organizational aspects of the biobank as a data collector, characteristics of the collected data, and maintaining control over personal data through informed consent.

Our results suggest that the medical scientific context is a key factor that influences the development of trust in biobanks and researchers in two ways. First, medical scientific research was generally perceived as trustworthy by participants and nonparticipants because it was considered beneficial for society, with several respondents perceiving a population-based biobank as a charitable organization. In other words, the underlying purpose of the scientific medical field contributes to the trustworthiness of research organizations. Previous studies have shown that diminution of trust in providing data can be related to commercialism (Critchley et al., 2015; Nicol et al., 2016; Steinsbekk, Myskja, & Solberg, 2013). Our results showed that this was indeed one of the main concerns about participation, especially relating to its withdrawal.

Second, the context of medical sciences was perceived as trustworthy, entailing qualified organizations and representatives. This perception reveals the importance of standardizing and harmonizing research systems based on clear principles, rules, and codes of conduct. It is nevertheless important to acknowledge that the contexts of medical science in other countries may be quite different and may not elicit similar degrees of trust (Sheikh & Hoeyer, 2018). Moreover, this factor has a bearing on the interpretation of the value placed by interviewees on informed consent. Our findings indicate that despite lack of recall of its content, informed consent was perceived by ex-participants, participants, and nonparticipants as a symbolic act to which a high moral value was assigned (Allen & McNamara, 2011).

The level of the interviewees' trust in centralized largescale data repository or biobanking was also influenced by the organizational aspects of the biobank as the data collector. The large scale of data repository and the sound reputation or trustworthiness of the responsible organization motivated participants and some nonparticipants to provide personal data despite their lack of knowledge regarding current security measures. Although providing data for Lifelines may entail risks concerning individuals' privacy and autonomy, most participants and nonparticipants experienced a sense of safety relating to the number of other participants in the biobank. By contrast, the large scale of data repository elicited a higher degree of distrust among ex-participants and several other nonparticipants. Previous findings indicate that the perception of strength and safety in numbers accounts for one of the reduced risks and increased positive outcomes (Park & Hinsz, 2006). This finding of difference in trust is in line with that of previous studies, suggesting that differences between participants and nonparticipants may be attributed to predisposed differences in terms of their concerns or generalizations of trust. Such differences have a bearing on information collection procedures as well as on the acceptance of new methods for collecting data (Nobile et al., 2016).

The trustworthiness of the biobank as an organization was evaluated by the interviewees partly on the basis of their perceptions of the motives, behaviors, and competence demonstrated by employees or other representatives of the biobank. This finding endorses previous findings that interpersonal relationships with health care representatives, such as general practitioners or biobank employees, matter in people's decisions to participate in research (Johnsson, 2013). The behaviors and communication styles of employees or other representatives of biobanks constitute the missing link, accounting for diminishing trust, as demonstrated by the ex-participants. This finding may indicate that trust in large-scale, centralized data repository is linked to perceptions of the trustworthiness of familiar entities (O'Neill, 2018), such that the aims and competencies of representatives are equated with the claims of the organization. Employees and representatives seem to be the gatekeepers of the bridge that stretches from trust, at one end of the spectrum, to distrust at the other. A lack of professionalism on their part is the final factor contributing to diminution of trust and the decision of individuals to withdraw their participation.

Ex-participants, nonparticipants, and even some participants had reservations about the process of combining information and data acquired from different contexts. The finding of the importance of the characteristics of the data, that is, their sensitivity and relevance to the research topic, highlights the significance of the context when collecting data or linking data sets. This context

is highly pertinent when considering the willingness of individuals to provide personal data. This challenges trends of centralizing data repositories in human research by internal and external linkage to other data sets, in particular to data sets about nonmedical contexts. Trust in a particular context and within organizations is contingent on the aims and norms that are valid in one context and is, therefore, of limited effectiveness in reducing complexity when attempting to optimize large-scale, centralized data repositories. These findings are in line with scholastic views on privacy that consider context as a key factor underlying concern (Barocas & Nissenbaum, 2014). Moreover, they support the bioethics guidelines formulated in the United Kingdom that state that the sensitivity of data is contingent on the context in which they are used and related to other data (Nuffield Council on Bioethics, 2015).

The findings of this study confirmed that “blind trust” on the part of participants in human research organizations was induced by the “halo effects” of the responsible organization and context (Hall et al., 2001). All (non)participants generalized their trust or distrust in the biobank and its context relating to data linkages to simplify the decision of contributing personal data. The level of risk associated, for example, with data optimization through linkages of data in a new setting can be estimated based on the experiences of interviewees combined with the perceived norms and values of the research organization or context. The role of trust therefore assumes prominence in decisions on whether or not to participate in research as a strategy for reducing complexity (Luhmann, 1979), particularly when the nature of data repository becomes increasingly complex.

In general, the perceptions of ex-participants, participants, and nonparticipants on data sharing varied, with most participants being generally positive about large-scale, centralized data repositories, focusing on opportunities that they offer rather than on threats. In contrast, nonparticipants and ex-participants had more concerns relating to these collections and offered reasons for these concerns. The findings of studies conducted among potential participants—4,659 and 15,650—in the United States and Europe, respectively, revealed similar attitudes (Gaskell et al., 2013; Kaufman et al., 2009). Although 90% of potential participants within a large cohort study in the United States expressed concerns about privacy, 60% of them stated that they would still participate in the study. The findings of the European study revealed that privacy concerns were less relevant to the intention to participate because of trust in key actors, for example, governments and universities. Therefore, the extension of existing trust seems to be an automatic process, whereas establishing trust from scratch could entail a cognitive rational process. This finding indicates that trust and distrust are distinct concepts in terms of structure. However, they are

equal in terms of their functions and therefore feature in decisions about risk taking (Luhmann, 1979). This finding is in line with those of previous studies suggesting that participants and nonparticipants differ in intrapersonal characteristics, such as generalized trust and distrust, which have relevance for information collection procedures as well for accepting new data repository methods, such as big data (Nobile et al., 2016).

In addition to highlighting relevant factors of (non)participation and trust in large-scale medical studies, our findings raise some ethical questions about current responsibilities and mechanisms for protecting individuals' integrity. The implementation of principles such as informed consent is highly valued in human research—apart from their being legally required—to protect participants' integrity (Eyal, 2014; Tassé & Kirby, 2017), and the implicit assumption underlying these principles is that participants are rational individuals. Our results revealed that when individuals need to decide whether or not to provide information to a biobank, feelings of trust in general, the research context, and the reputation of the organization seem to be more relevant than knowledge about potential threats to their personal integrity. These factors may be even more relevant to decisions relating to the acceptance of centralized data repository and data linkages.

The above discussion foregrounds the ethical issue of self-determination through one instance of detailed informed consent conceived as the basis of the consent and information process, particularly in relation to biobanks, which are public entities. When biobanks are perceived as charitable organizations, the societal context appears to be a prominent factor in the decision to entrust an organization with personal data. However, the current consent and information-gathering procedures mainly focus on risks for individuals that are of a physical nature (Broekstra et al., 2017; D'Abramo, 2015). In light of these issues, our results suggest that current procedures, such as broad informed consent obtained in large cohort studies and biobanks (Steinsbekk, Myskja, & Solberg, 2013), are inadequate for safeguarding participants' personal integrity and trust. On one hand, the societal context and potential implications of participation, such as reduced privacy, are not completely addressed, and on the other hand, participants are not protected against psychological biases at an individual level.

Best Practices

The combination of our findings leads to several recommendations for the management of biorepositories and data repositories. First, the findings argue for an adoption of different measures in information and consent procedures, such as models of governance or consent based on solidarity to ensure partic-

ipants' autonomy (Mittelstadt & Floridi, 2016; Prainsack & Buyx, 2013). These models can accommodate risks on individual as well as societal levels, given the focus on decreasing potential threats to individual participants versus providing more complete and relevant information for research or other public goods. Second, it is important to design continuous communication strategies for participants and the public while taking into consideration organizational aspects and characteristics of collected data that influence levels of trust, such as scale of organization, quality of employees, or context of collected data. This finding highlights the moral responsibility of the scientific community to engage in public outreach regarding large-scale, centralized data repository in scientific human research, especially in the medical sciences.

One strength of our study, which entailed the conduct of in-depth semi-structured interviews, is that it has shed some light on the complex decision-making process for providing personal data to a biobank. To the best of our knowledge, this study is the first to identify and compare factors associated with the trust of ex-participants, participants, and nonparticipants in large-scale, centralized data repository for human research. Moreover, through its focus not only on members of the public as nonparticipants in data repository but also on ex-participants, the study has provided a broader perspective on decision making that pertains to the provision of personal data in large-scale, centralized data repository.

The study revealed differences in the perceptions, concerns, and elements entailed in trust building associated with the sharing of data among these three groups. This comparison of these three groups is, however, tentative, as the sample size was small, especially for the group of ex-participants. Not surprisingly, most ex-participants who were approached, as well as their partners, were unwilling to participate in our study. The fact that the biobank itself recruited the ex-participants probably decreased their willingness to participate in our study. Importantly, saturation occurred during the interviews held with participants and nonparticipants as well as with ex-participants, suggesting that our study results were not strongly affected by issues of representativeness for any of the three groups.

Research Agenda

The present study provides some initial evidence relating to the factors that determine individuals' trust levels and their decisions to withdraw from participation in human research in the field of medical science. Our results point to the relevance of several factors, for example, the importance of capable and trustworthy representatives and the reputation of an organization. Evidently, further

research on this topic is needed. For example, it seems worthwhile to focus on the potential role of individual characteristics entailed in generalized trust and the decision to share personal data with biobanks, especially via data linkages. These factors could be investigated more effectively using quantitative methods.

Furthermore, the results of our study suggest that knowledge about risks, safety measures, and the benefits of technologies could raise concerns of privacy and security among some nonparticipants and ex-participants, especially in contexts where trust levels in data repository procedures are low. Proactive communication and information-gathering procedures are therefore critical for avoiding distrust. Two previous studies showed that an inverse relationship exists between perceived risks and perceived benefits, which determined the acceptance of technologies (Siegrist, 2000; Siegrist & Cvetkovich, 2000). Such a relationship was not, however, found for individuals who reported having knowledge about these risks and benefits (Siegrist & Cvetkovich, 2000). The conduct of more such quasi-experimental studies would contribute to an understanding of the effects of such interventions entailing communication and information-gathering procedures. Furthermore, studies should focus on differences among domains of human research to determine which elements are decisive for trust building. Given that we only focused on the medical field, a comparison of different contexts would be pertinent.

Educational Implications

These findings suggest several courses of action for employees from data repositories in human research, and the researchers themselves. First, trust is enhanced by qualified and trustworthy employees, which may have training and educational implications for biobanks as public goods seeking a large number of participants. Moreover, there is an important role for governments in providing guarantees of the commitments and competencies of institutes and representatives within research systems via training and education (Gilson, 2003). Second, proactive communication is key for preventing diminution of trust, in particular for changes in data collection. Trust and engagement can be enhanced through face-to-face contact (Gaskell et al., 2013). Ex-participants reported that the organization or its representatives did not communicate proactively with them. Prevention of withdrawal of participation could, therefore, be done by being able to provide the right explanation on time, for example, about data security, organizational aspects, or limits of context of data repository.

Conclusion

In conclusion, the findings of the present study indicate that the decision to provide personal data for medical scientific research and to accept centralized data repository or data linkages is not one that is made easily. It relies heavily on generalized trust, which is built upon factors relating to the context of data repository, but it is also contingent on individuals' attitudes and values. Recent proposals relating to governance in biobanking (Mittelstadt & Floridi, 2016; Prainsack & Buyx, 2013) and informed consent (BudinLjønsne et al., 2017) are heading in the direction of customization and greater transparency rather than focusing on individual participants and their needs. Contexts perceived as trustworthy, dynamic data governance, and informed consent models and ensuring benefits for the public from research may enhance potential participants' willingness to provide personal data. These concepts foreground reciprocity and interactions between researchers and (non)participants, with the aim of tailoring research within a wider context, thereby contributing to the building of trust in human research. However, challenges lie ahead for achieving the sustainable use of centralized large-scale data repositories in human research.

References

- Allen, J., & Mcnamara, B. (2011). Reconsidering the value of consent in biobank research. *Bioethics*, 25(3), 155-166. <https://doi.org/10.1111/j.1467-8519.2009.01749.x>
- Baker, S. E., & Edwards, R. (2012). *How many qualitative interviews is enough?* [National Centre for Research Methods Review Paper]. http://eprints.ncrm.ac.uk/2273/4/how_many_interviews.pdf
- Balliet, D., & Van Lange, P. A. M. (2013). Trust, conflict, and cooperation: A meta-analysis. *Psychological Bulletin*, 139(5), 1090-1112. <https://doi.org/10.1037/a0030939>
- Barocas, S., & Nissenbaum, H. (2014). Big data's end run around anonymity and consent. In L. Julia, V. Stodden, B. Stefan, & H. Nissenbaum (Eds.), *Privacy, big data and the public good* (pp. 44-75). Cambridge University Press. <https://doi.org/http://dx.doi.org/10.1017/CBO9781107590205>
- Bhattacharya, R., Devinney, T. M., & Pillutla, M. M. (1998). A formal model of trust based on outcomes. *Academy of Management Review*, 23(3), 459-472. <https://doi.org/10.5465/AMR.1998.926621>
- Boyd, d., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662-679. <https://doi.org/10.1080/1369118X.2012.678878>
- Broekstra, R., Maeckelberghe, E. L. M., & Stolk, R. P. (2017). Written informed consent in health research is outdated. *European Journal of Public Health*, 27(2), 194-195. <https://doi.org/10.1093/eurpub/ckw198>
- Budin-Ljøsne, I., Teare, H. J. A., Kaye, J., Beck, S., Bentzen, H. B., Caenazzo, L., Collett, C., D'Abramo, F., Felzmann, H., Finlay, T., Javaid, M. K., Jones, E., Katić, V., Simpson, A., & Mascalzoni, D. (2017). Dynamic consent: A potential solution to some of the challenges of modern biomedical research. *BMC Medical Ethics*, 18(1), 4. <https://doi.org/10.1186/s12910-016-0162-9>
- Coyne, I. T. (1997). Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries? *Journal of Advanced Nursing*, 26(3), 623-630. <http://doi.org/10.1046/j.1365-2648.1997.t01-25-00999.x>
- Critchley, C., Nicol, D., & McWhirter, R. (2016). Identifying public expectations of genetic biobanks. *Public Understanding of Science*, 26(6), 671-687. <https://doi.org/10.1177/0963662515623925>
- Critchley, C., Nicol, D., & Otlowski, M. (2015). The impact of commercialisation and genetic data sharing arrangements on public trust and the intention to participate in biobank research. *Public Health Genomics*, 18(3), 160-172. <https://doi.org/10.1159/000375441>

- Critchley, C., Nicol, D., Otlowski, M., & Stranger, M. (2012). Predicting intention to biobank: A national survey. *European Journal of Public Health*, 22(1), 139-144. <https://doi.org/10.1093/eurpub/ckq136>
- D'Abramo, F. (2015). Biobank research, informed consent and society. Towards a new alliance? *Journal of Epidemiology and Community Health*, 69(11), 1125-1128. <https://doi.org/10.1136/jech-2014-205215>
- Eurobarometer. (2014). *Public perceptions of science, research and innovation* (Special Eurobarometer 419). <https://doi.org/10.2777/95599>
- Eyal, N. (2014). Using informed consent to save trust. *Journal of Medical Ethics*, 40(7), 437-444. <https://doi.org/10.1136/medethics-2012-100490>
- Ferrin, D. L., Dirks, K. T., & Shah, P. P. (2006). Direct and indirect effects of third-party relationships on interpersonal trust. *Journal of Applied Psychology*, 91(4), 870-883. <https://doi.org/10.1037/0021-9010.91.4.870>
- Friese, S. (2013). *ATLAS.ti 7 user guide and reference*. ATLAS.ti Scientific Software Development GmbH.
- Gaskell, G., Gottweis, H., Starkbaum, J., Gerber, M. M., Broerse, J., Gottweis, U., Hobbs, A., Helén, I., Paschou, M., Snell, K., & Soulier, A. (2013). Publics and biobanks: Pan-European diversity and the challenge of responsible innovation. *European Journal of Human Genetics*, 21(1), 14-20. <https://doi.org/10.1038/ejhg.2012.104>
- Gibson, S. G., Axler, R. E., & Lemmens, T. (2017). Transparency of biobank access in Canada: An assessment of industry access and the availability of information on access policies and resulting research. *Journal of Empirical Research on Human Research Ethics*, 12(5), 310-325. <https://doi.org/10.1177/1556264617723137>
- Gilson, L. (2003). Trust and the development of health care as a social institution. *Social Science & Medicine*, 56(7), 1453-1468. [https://doi.org/10.1016/S0277-9536\(02\)00142-9](https://doi.org/10.1016/S0277-9536(02)00142-9)
- Graeff, T. R., & Harmon, S. (2002). Collecting and using personal data: Consumers' awareness and concerns. *Journal of Consumer Marketing*, 19(4), 302-318. <https://doi.org/10.1108/07363760210433627>
- Hall, M. A., Dugan, E., Zheng, B., & Mishra, A. K. (2001). Trust in physicians and medical institutions: What is it, can it be measured, and does it matter? *The Milbank Quarterly*, 79(4), 613-639. <http://www.jstor.org/stable/3350617>
- Hansson, M. G. (2005). Building on relationships of trust in biobank research. *Journal of Medical Ethics*, 31(7), 415-418. <https://doi.org/10.1136/jme.2004.009456>
- Johnsson, L. (2013). *Trust in biobank research: Meaning and moral significance*. Uppsala Universitet. <http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-192295>
- Kaufman, D. J., Murphy-Bollinger, J., Scott, J., & Hudson, K. L. (2009). Public opin-

- ion about the importance of privacy in biobank research. *American Journal of Human Genetics*, 85(5), 643-654. <https://doi.org/10.1016/j.ajhg.2009.10.002>
- Kelley, M., James, C., Alessi Kraft, S., Korngiebel, D., Wijangco, I., Rosenthal, E., Joffe, S., Cho, M. K., Wilfond, B., & Lee, S.S. (2015). Patient perspectives on the learning health system: The importance of trust and shared decision making. *The American Journal of Bioethics*, 15(9), 4-17. <https://doi.org/10.1080/15265161.2015.1062163>
- Kettis-Lindblad, Å., Ring, L., Viberth, E., & Hansson, M. G. (2006). Genetic research and donation of tissue samples to biobanks. What do potential sample donors in the Swedish general public think? *European Journal of Public Health*, 16(4), 433-440. <https://doi.org/10.1093/eurpub/ckl198>
- Keymolen, E. (2016). *Trust on the line: A philosophical exploration of trust in the networked era*. Erasmus University Rotterdam. <http://hdl.handle.net/1765/93210>
- Krokstad, S., Langhammer, A., Hveem, K., Holmen, T. L., Midthjell, K., Stene, T. R., Bratberg, G., Heggland, J., & Holmen, J. (2013). Cohort profile: The HUNT Study, Norway. *International Journal of Epidemiology*, 42(4), 968-977. <https://doi.org/10.1093/ije/dys095>
- Kvale, S. (2007). *Doing interviews: The SAGE qualitative research kit*. SAGE. <https://doi.org/http://dx.doi.org/10.4135/9781849208963>
- Layder, D. (1998). Sociological practice: Linking theory and social research.
- Lemke, A. A., Wolf, W. A., Hebert-Beirne, J., & Smith, M. E. (2010). Public and biobank participant attitudes toward genetic research participation and data sharing. *Public Health Genomics*, 13(6), 368-377. <https://doi.org/10.1159/000276767>
- Lewicki, R. J., McAllister, D. J., & Bies, R. I. (1998). Trust and distrust: New relationships and realities. *Academy of Management Review*, 23(3), 438-458. <https://doi.org/10.2307/259288>
- Luhmann, N. (1979). *Trust and power*. John Wiley. <https://books.google.nl/books?id=WZu4AAAAIAAJ>
- Mason, M. (2010). Sample size and saturation in PhD studies using qualitative interviews. *Forum Qualitative Sozialforschung/ Forum: Qualitative Social Research*, 11(3), Article 8. <http://nbn-resolving.de/urn:nbn:de:0114-fqs100387>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709-734. <https://doi.org/10.2307/258792>
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1-21. <https://doi.org/10.1177/2053395171667969>

- Mittelstadt, B. D., & Floridi, L. (2016). The ethics of big data: Current and foreseeable issues in biomedical contexts. *Science and Engineering Ethics* 22(2), 303-341. <https://doi.org/10.1007/s11948-015-9652-2>
- Nicol, D., Critchley, C., McWhirter, R., & Whitton, T. (2016). Understanding public reactions to commercialization of biobanks and use of biobank resources. *Social Science & Medicine*, 162, 79-87. <https://doi.org/10.1016/j.socscimed.2016.06.028>
- Nobile, H., Bergmann, M. M., Moldenhauer, J., & Borry, P. (2016). Participants' accounts on their decision to join a cohort study with an attached biobank: A qualitative content analysis study within two German studies. *Journal of Empirical Research on Human Research Ethics*, 11(3), 237-249. <https://doi.org/10.1177/1556264616657463>
- Nobile, H., Vermeulen, E., Thys, K., Bergmann, M. M., & Borry, P. (2013). Why do participants enroll in population biobank studies? A systematic literature review. *Expert Review of Molecular Diagnostics*, 13(1), 35-47. <https://doi.org/10.1586/erm.12.116>
- Nuffield Council on Bioethics. (2015). The collection, linking and use of data in biomedical research and health care: Ethical issues, 198.
- O'Neill, O. (2018). Linking trust to trustworthiness. *International Journal of Philosophical Studies*, 26(2), 293-300. <https://doi.org/10.1080/09672559.2018.1454637>
- Park, E. S., & Hinsz, V. B. (2006). "Strength and safety in numbers": A theoretical perspective on group influences on approach and avoidance motivation. *Motivation and Emotion*, 30(2), 135-142. <https://doi.org/10.1007/s11031-006-9024-y>
- Parks, C. D. (2015). Determinants of cooperation in social dilemmas. In D. A. Schroeder & W. G. Graziano (Eds.), *The Oxford handbook of prosocial behavior*. <https://doi.org/10.1093/oxfordhb/9780195399813.013.005>
- Parks, C. D., Joireman, J., & Van Lange, P. A. M. (2013). Cooperation, trust, and antagonism: How public goods are promoted. *Psychological Science in the Public Interest*, 14(3), 119-165. <https://doi.org/10.1177/1529100612474436>
- Petersen, I., Desmedt, C., Harris, A., Buffa, F., & Kollek, R. (2014). Informed consent, biobank research, and locality: Perceptions of breast cancer patients in three European countries. *Journal of Empirical Research on Human Research Ethics*, 9(3), 485-5. <https://doi.org/10.1177/1556264614540600>
- Prainsack, B., & Buyx, A. (2013). A solidarity-based approach to the governance of research biobanks. *Medical Law Review*, 21(1), 71-91. <https://doi.org/10.1093/medlaw/fws040>
- Rahm, A. K., Wrenn, M., Carroll, N. M., & Feigelson, H. S. (2013). Biobanking for

- research: A survey of patient population attitudes and understanding. *Journal of Community Genetics*, 4(4), 445-450. <https://doi.org/10.1007/s12687-013-0146-0>
- Ridgeway, J. L., Han, L. C., Olson, J. E., Lackore, K. A., Koenig, B. A., Beebe, T. J., & Ziegenfuss, J. Y. (2013). Potential bias in the bank: What distinguishes refusers, nonresponders and participants in a clinic-based biobank? *Public Health Genomics*, 16(3), 118-126. <https://doi.org/10.1159/000349924>
- Rivera, S. M., Brothers, K. B., Cadigan, R. J., Harrell, H. L., Rothstein, M. A., Sharp, R. R., & Goldenberg, A. J. (2017). Modernizing research regulations is not enough: It's time to think outside the regulatory box. *American Journal of Bioethics*, 17(7), 13. <https://doi.org/10.1080/15265161.2017.1328899>
- Robinson, J. O., Slashinski, M. J., Wang, T., Hilsenbeck, S. G., & McGuire, A. L. (2013). Participants' recall and understanding of genomic research and large-scale data sharing. *Journal of Empirical Research on Human Research Ethics*, 8(4), 42-52. <https://doi.org/10.1525/jer.2013.8.4.42>
- Saunders, M. N. K., Dietz, G., & Thornhill, A. (2014). Trust and distrust: Polar opposites, or independent but coexisting? *Human Relations*, 67(6), 639-665. <https://doi.org/10.1177/0018726713500831>
- Scholtens, S., Smidt, N., Swertz, M. A., Bakker, S. J. L., Dotinga, A., Vonk, J. M., van Dijk, F., van Zon, S. K., Wijmenga, C., Wolffenbuttel, B. H., & Stolk, R. P. (2015). Cohort profile: LifeLines, a three-generation cohort study and biobank. *International Journal of Epidemiology*, 44(4), 1172-1180. <https://doi.org/10.1093/ije/dyu229>
- Sheikh, Z., & Hoeyer, K. (2018). "That is why I have trust": Unpacking what "trust" means to participants in international genetic research in Pakistan and Denmark. *Medicine, Health Care and Philosophy*, 21(2), 169-179. <https://doi.org/10.1007/s11019-017-9795-9>

- Siegrist, M. (2000). The influence of trust and perceptions of risks and benefits on the acceptance of gene technology. *Risk Analysis*, 20(2), 195-203. <https://doi.org/10.1111/0272-4332.202020>
- Siegrist, M., & Cvetkovich, G. (2000). Perception of hazards: The role of social trust and knowledge. *Risk Analysis*, 20(5), 713719. <https://doi.org/10.1111/0272-4332.205064>
- Sorani, M. D., Yue, J. K., Sharma, S., Manley, G. T., & Ferguson, A. R., Cooper, S. R., Dams O'Connor, K., Gordon, W. A., Lingsma, H. F., Maas, A. I. R., Menon, D. K., Morabito, D. J., Mukherjee, P., Okonkwo, D. O., Puccio, A. M., Valadka, A. B., & Yuh, E. L. (2015). Genetic data sharing and privacy. *Neuroinformatics*, 13(1), 16. <https://doi.org/10.1007/s12021-014-9248-z>
- Steinsbekk, K. S., Myskja, B. K., & Solberg, B. (2013). Broad consent versus dynamic consent in biobank research: Is passive participation an ethical problem? *European Journal of Human Genetics*, 21(9), 897-902. <https://doi.org/10.1038/ejhg.2012.282>
- Steinsbekk, K. S., Ursin, L. Ø., Skolbekken, J.-A., & Solberg, B. (2013). We're not in it for the money-lay people's moral intuitions on commercial use of "their" biobank. *Medicine, Health Care and Philosophy*, 16(2), 151-162. <https://doi.org/10.1007/s11019-011-9353-9>
- Stricker, B. H. (2017). Epidemiology and "big data." *European Journal of Epidemiology*, 32(7), 535-536. <https://doi.org/10.1007/s10654-017-0294-3>
- Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., & Collins, R. (2015). UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLOS Medicine*, 12(3), Article e1001779. <https://doi.org/10.1371/journal.pmed.1001779>

- Tassé, A. M., & Kirby, E. (2017). Is written informed consent outdated? *European Journal of Public Health*, 27(2), 194-195. <https://doi.org/10.1093/eurpub/ckw198>
- UK Biobank Coordinating Centre. (2007). UK biobank: Protocol for a large-scale prospective epidemiological resource UK Biobank Coordinating Centre Stockport. *Design*, 6, 1-112.
- van Staa, T.-P., Goldacre, B., Buchan, I., & Smeeth, L. (2016). Big health data: The need to earn public trust. *British Medical Journal*, 354, Article i3636. <https://doi.org/10.1136/bmj.i3636>
- Ziebland, S., & McPherson, A. (2006). Making sense of qualitative data analysis: An introduction with illustrations from DIPEX (personal experiences of health and illness). *Medical Education*, 40(5), 405-414. <https://doi.org/10.1111/j.1365-2929.2006.02467.x>

Chapter 4

Demographic and prosocial intrapersonal characteristics of biobank participants and refusers: the findings of a survey in the Netherlands

Abstract

Research in genetics relies heavily on voluntary contributions of personal data. We aimed to acquire insights into the differences between participants and refusers of participation in a Dutch population-based biobank. Accordingly, we assessed the demographic and prosocial intrapersonal characteristics of respondents who participated ($n = 2615$) or refused to participate ($n = 404$) in the Lifelines biobank and databank. Our results indicated that health-related values critically influence participation decisions. The participation threshold for Lifelines was determined by an absence of health-related values and of trust in government. Therefore, considering these factors in communication and recruitment strategies could enhance participation in biomedical research. No indications were found of a stronger general prosociality of participants or their trust in researchers beyond the context of biobanking. This emphasizes the contextual understanding of the decision of participation in biobanking. Our findings may contribute to improving recruitment strategies by incorporating relevant values and/or highlighting prosocial benefits. Moreover, they foreground the need to address trust issues in collaborations between data repositories and commercial companies. Future research should explore how prosocial intrapersonal characteristics drive participation and withdrawal decisions and relate to contextual attributes.

Keywords: biobanks; research participation; refusal ; values ; trust

This chapter is previously published as: Broekstra, R., Aris-Meijer, J., Maeckelberghe, E.L.M, Stolk, R.P., Otten S. Demographic and prosocial intrapersonal characteristics of biobank participants and refusers: the findings of a survey in the Netherlands. Eur J Hum Genet 29, 11–19 (2021). <https://doi.org/10.1038/s41431-020-0701-1>

Introduction

Genetic and biomedical research rely heavily on voluntary contributions of personal data [1]. Simultaneously, an increase in calls for data has led to mounting concerns about data security and privacy [2]. These concerns might lead to lower willingness to participate in biomedical research [3–6]. Therefore, a study of the differences between those who accept and decline invitations to participate in a biobank could yield valuable insights into effective strategies for recruiting participants [7].

Studies have identified various factors that indicate whether people are willing to participate in biobanks or not. There is ample evidence that several demographic characteristics are associated with the intention to participate in biobanks, e.g., being highly educated [8–11], having a partner [6, 12], being non-religious [10, 13] and having better (self-reported) health [8, 14]. Other demographic characteristics are more inconsistently associated, such as being older [3, 8, 15] or younger [10, 11] vs. no age-effect [9, 16, 17]; being male [3, 17] or female [16] vs. no gender effect [9, 10].

Psychological characteristics, such as trust in and concern for others, provide a stable pattern of associations with willingness to participate. Studies report that participation willingness in biobanks is strongly associated with having greater societal trust [3, 8], with higher levels of social engagement [18], and with being registered as organ or blood donor [8]. Societal trust refers to a generalized trust in the government as system or institution [19], and in unknown other citizens [20]. Other strong associations with participation have been found for factors specific to the biobanking domain, such as trust in biobanks [6, 9] or positive attitudes towards biobanking in general [3, 21, 22], and towards biomedical research [21, 22].

Attitudes, values and behaviours reflecting cooperative tendencies and confidence in cooperation—which we here summarize as prosocial intrapersonal characteristics—independently contributed to willingness to participate in biobanking [9, 21]. The key role of these prosocial characteristics is in line with studies emphasizing the importance of prosocial motives for participation in biobanking [3, 4, 23]. Moreover, similar associations were found for prosocial characteristics with decisions about cooperation in other societal contexts, for example, participation in voluntary work [24] or in smart energy systems [25]. Especially, societal trust and domain-specific values, e.g. concern for environment, proved to be robust predictors in these contexts [26–28].

Furthermore, psychological characteristics seem to explain non-participation, and in particular refusal, better than demographic characteristics. For

example, a study in the USA found that women and older patients were more likely to respond to recruitment, but these demographics were not associated with refusal. Instead, people refusing participation in biobanking looked, just as participants, more often and more actively for information about biobanking than non-participants not responding to recruitment request, while participants as well as both non-participant groups had the same prior knowledge [16].

It is possible that thresholds for participation arise due to competing values about society or concerns about biobanking. A recent study in Finland showed that donating blood for biobanking purposes was perceived with more concerns than donating blood for treatment of patients [29]. Trust reduced complexity of decisions about participation by perceiving fewer risks [19, 30]. A lack of trust indicates an increased anxiety about harm [31]. This anxiety may be exacerbated when sharing data with biobanks, in particular involving commercial enterprises or researchers. Studies showed that refusal in biobanking is related to societal and individual concerns, such as fear of discovering possible genetic predisposition to certain diseases leading to social stigma [21]; worry that information may be used against personal interest resulting into discrimination [16, 21, 30]; or concerns about commercial purposes eventually derogating public goods [21, 23, 30].

Together, the evidence summarized above suggests that psychological characteristics could be relevant predictors of both participation and refusal in biobanking. We investigated people who either participated or explicitly refused to participate in a Dutch population-based biobank. We aimed to contribute insights into (non-)participation in biobanking by simultaneously investigating demographic and prosocial intrapersonal characteristics to elucidate their independent roles in the decision to participate or refuse participation, regardless of the motives for these decisions.

Materials and methods

Procedure and participants

We conducted an online survey of a sample of the general population in the Northern provinces of the Netherlands ($n = 3019$), comprising 2615 participants in a Dutch biobank, and 404 citizens who refused or signaled refusal to contribute to Lifelines or similar large-scale centralized data repository for scientific medical research. Our sample of participants was recruited via Lifelines in May 2018. The response rate for the invitation of this additional study was 22.2% without incentives. Lifelines recruited its participants between 2006 and 2013 from the general population in the northern provinces of the Netherlands, and providing

a representative sample ($n = 167,000$) [12, 32]. Our sample of refusers was recruited among a representative respondents' panel of the Northern Dutch population via the DirectResearch EUpanel ($n = 30,600$) in August of the 2018. These respondents received a small financial incentive for their studies, resulting in a response rate of 41.4%, respectively. Both organizations applied the same recruitment protocol and stratified the samples for sex and age. The subsample of biobank participants was provided by Lifelines themselves, since they exclusively recruit data from their participants. These participants consented between 2006 and 2013 to participate in a longitudinal prospective cohort study of onset, risk factors and course of chronic diseases [32]. Refusers stemmed from the DirectResearch sample from the northern general population. Self-reported refusal was measured via two items at the beginning of our questionnaire. First, these were people who indicated that they did not participate "in large scale medical scientific studies, for example Lifelines" ($n = 1738$). Next, within this broader sample of non-participants, we identified refusers via their affirmation of the statement "I did not want/would not want to participate in this type of research" ($n = 404$). The mean age of Lifelines' respondents was 56 ($SD = 15.88$) with 50.5% being male, whereas the mean age of DirectResearch' respondents after exclusion was 56 ($SD = 14.18$) with 48.5% being male (Table 1 and Supplementary Table 1).

Table 1. Demographic characteristics and self-reported health of participants and refusers of a Dutch population-based biobank

	Participants		Refusers		P
Age	56.27	-15.88	56.09	-14.18	0.83
Gender					0.47
Male	1291	-50.50%	179	-48.50%	
Female	1264	-49.50%	190	-51.50%	
Marital status					≤ 0.01
Registered partner	2,078	-79.50%	269	-66.60%	
No registered partner	536	-20.50%	134	-33.40%	
Educational level					≤ 0.01
High	1,083	-41.40%	189	-46.80%	
Moderate	938	-35.90%	50	-37.10%	
Low	559	-21.70%	60	-14.90%	
Paid Job	1,460	-55.80%	205	-50.70%	0.07
Religious	972	-37.20%	140	-34.70%	0.48
Residence					0.04
Rural	2,019	-77.20%	295	-73.00%	
Urban	537	-20.50%	102	-25.20%	
Self-reported health (1= very poor, 5 = very good)	3.97	-0.71	3.71	-0.74	≤ 0.01

Note: The values shown in the table are either means with standard deviations or numbers (n) with percentages (%) relating to the total sample or mean values with standard deviations.

Measures

Demographic characteristics and self-reported health

We measured previously identified demographic characteristics associated with participation or willingness to participate in a biobank, such as marital status [6, 12], education level [8–11], religion [10, 13], residential areas [12], and self-reported general health [8, 14]. Work status was included as additional indicator of socioeconomic status.

Prosocial intrapersonal characteristics

Prosocial orientation and values.

We applied two measures to determine respondents' prosocial orientations. The first was their actual prosocial behaviour reflected in their organ and blood donor status and the frequency of their charitable donations. The second was the degree to which respondents cared about others' outcomes in resource allocation choices, which was measured using six key items of the Social Value Orientation (SVO) scale. This scale is designed to measure the magnitude of individuals' general concern for others [33]. Replicating previous research, our results indicated excellent reliability with $\alpha > 0.90$ [34].

We used eleven items based on the theory of basic values to assess motivations of self-interest and beyond [35]. Moreover, we applied an adapted version of the Environmental Personal Values Questionnaire (E-PVQ) scale [26]. Although our focus primarily was on behaviour in the context of healthcare, we investigated values related to a wider context, e.g., about the environment. This might distinguish individuals acting based on their general concern for others, themselves and a common environment. In addition to the original subscales measuring hedonic values (individual pleasure, e.g., "it is important to have fun") and subscale on biospheric values (concern about the environment, e.g., "it is important to protect the environment"), we included a subscale on healthspheric values (health and healthcare concerns, e.g. "it is important to live a fit life"). The reliability was good with $\alpha > 0.80$ on all subscales, but they had a strong intercorrelation (Table 2).

Societal trust.

Societal trust was measured based on the framework proposed by Mayer et al. [36]. We used six items reflecting trust in the government and trust in other citizens on a fivepoint Likert-scale, e.g., "the government can take care of her citizens"; "the government acts with good intentions"; "people are trustworthy". The six items tap into various aspects of societal trust, namely, competence, good intentions, and trustworthiness [36]. The combined scale had a

Table 2. Correlation Table

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. Age	1											
2. Educational level (1 = low, 2 = moderate, 3 = high)	-0.21	1										
3. Self-reported health	-0.07	0.15	1									
4. Frequency of charitable donations	0.16	0.06	.05 ^s	1								
5. Social Value Orientation score	0.01	0.06	0.02	0.15	1							
6. Healthspenic values	0.17	-0.02	0.26	0.14	0.12	1						
7. Hedonic values	-0.07	-0.05	0.16	0.07	0.09	0.71	1					
8. Biospheric values	0.23	-0.01	0.12	0.18	0.14	0.73	0.54	1				
9. Societal trust	0.02	0.22	0.22	0.18	0.13	0.2	0.13	0.17	1			
10. TDM&H/government	-0.4 ^s	0.11	0.12	0.11	0.12	0.19	0.19	0.16	0.42	1		
11. TDM&H/commercial enterprises	0.06	-0.14	0.03	0.07	0.05	0.05	0.01	0.02	0.24	0.34	1	
12. TDM&H/commercial researchers	.04 ^s	0.05	0.07	0.09	0.08	0.02	-0.01	.04 ^s	0.34	0.47	0.4	1

Note: **Bold** is significant with $p < 0.01$; **Bold^s** is significant with $p < 0.05$

good reliability ($\alpha = 0.86$).

In addition, five items measured trust in a domain specific way (i.e., referring to large-scale data repository), namely, trust in different types of research organizations their data management and handling (i.e., hospital, government, university, large-sized commercial enterprises, and medium-sized enterprises (SMEs)). For example, we used items “I believe that the hospital correctly, adequately and fairly store and treat my personal data”. Besides trust in organizations as entity, we measured trust in specific common research employees their data management and handling with four items (i.e., hospital researcher, university researcher, market researcher, polling researcher). For example, an item was “I trust employees of the hospital conducting research with a correct, fair and careful approach”. We conducted a principal component analysis with an oblimin rotation, due to several high intercorrelations ($r > 0.50$). The analysis showed that the items could be reduced to three factors explaining 71.90% of the total variance. The first factor, with four items, was labelled “Trust in data management and handling of government” (including researchers’ data practices within its institutes) and explained 45.74% of the total variance ($\alpha = 0.83$). The second factor, “Trust in data management and handling of commercial organizations”, comprised two items and explained 14.24% variance ($\alpha = 0.65$). The third factor, “Trust in data management and handling of commercial researchers”, comprised two items and explained 11.92% of the total variance ($\alpha = 0.47$). Table 2 shows that the correlations between components were low to moderate ($r < 0.40$).

Data analysis

We calculated descriptive statistics for demographic and psychological variables, applying the chi-square test or independent t-tests to calculate differences in means between participants and refusers where appropriate. We performed multivariate binary logistic regression analyses with a manual entering strategy to build respondents’ profiles based on their psychological characteristics while controlling for demographic characteristics. All analyses were conducted using SPSS, Version 25.0 [37].

Results

Demographic characteristics and self-reported health

Table 1 shows that more participants than refusers had registered partnerships including marriages (79.5 vs. 66.6%, $p < 0.01$). There was a positive though non-significant trend for having a job (55.8 vs. 50.7%, $p = 0.07$). While partici-

pants reported better general health status (3.97 vs. 3.71, $p < 0.01$), refusers had a higher educational level (46.8 vs. 41.4%, $p < 0.01$), were more commonly urban dwelling (77.2 vs. 73%, $p < 0.01$) and less religious (37.2 vs. 34.7%, $p < 0.01$).

Prosocial intrapersonal characteristics

Prosocial orientation and values.

Table 3 shows that compared to refusers, participants were more commonly contributors to charities (2.68 vs. 2.44, $p < 0.01$), blood donors (11.7 vs. 5.4%, $p < 0.01$) and organ donors (55.8 vs. 36.7%, $p < 0.01$). This finding reflects the stronger concerns about others among participants, as indicated by their SVO scores compared with refusers' scores (32.25 vs. 29.95, $p < 0.01$). Moreover, they appeared to be more concerned about health, a healthy lifestyle (3.99 vs. 3.75, $p < 0.01$), and their own pleasure (4.25 vs. 4.09, $p < 0.01$) than about their environment (3.76 vs. 3.83, $p < 0.01$).

Table 3. Prosocial Intrapersonal Characteristics of participants and refusers of the Lifelines biobank

	Participants		Refusers		P
Frequency of donations to charities	2.68	-1.08	2.44	-1.04	≤ 0.01
Blood donor	306	-11.70%	22	-5.40%	≤ 0.01
Organ donor	1,460	-55.80%	149	-36.70%	≤ 0.01
Social Value Orientation score	32.25	-14.19	29.94	-13.85	≤ 0.01
Values					
Healthspheric	3.99	-0.66	3.73	-0.73	≤ 0.01
Hedonic	4.25	-0.65	4.09	-0.7	≤ 0.01
Biospheric	3.76	-0.75	3.83	-0.75	≤ 0.01
Societal trust	3.25	-0.63	3.02	-0.74	≤ 0.01
TDM&H/government	3.82	-0.54	3.62	-0.71	≤ 0.01
TDM&H/commercial enterprises	2.73	-0.73	2.74	-0.76	0.81
TDM&H/commercial researchers	3.34	-0.72	3.44	-0.67	≤ 0.01

Note: The values in the table are either means with standard deviations or numbers (n) with percentages (%) relating to the total sample. TDM&H denotes Trust in Data Management and Handling.

Societal trust.

Participants had higher levels of trust in the government and in other citizens compared to refusers (3.25 vs. 3.14, $p < 0.01$), especially regarding the government's management and handling of data including researchers' data management practices within its institutes (3.82 vs. 3.62, $p < 0.01$). However, participants were less trusting of researchers' data handling and management practices within commercial enterprises (3.34 vs. 3.44, $p < 0.01$). There was no difference in levels of trust concerning commercial enterprises among participants and refusers (2.73 vs. 2.74, $p > 0.50$).

Main analysis

We started our main analysis by testing the assumptions of our logistic regression, which were all met. First, there were no indications of multicollinearity; all VIF scores were < 10.0 [38]. Second, some of the continuous variables showed a non-linear relationship of the independent variable to their log odds, as shown by the Box Tidwell test [39]. These variables were adjusted for their non-linear relationship as long as the model fit significantly improved (Table 4). Lastly, the sample size was considered sufficient applying the rule of thumb suggested by Peduz and collaborators of $n > 10 \times 14/0.1 = 1400$ and not finding expected cell counts < 5 [40].

The results of the multivariate logistic regression analyses of potential predictors of participation revealed that our model performed well with a significant goodness of fit ($\chi^2(26) = 615.414$, $p < 0.01$). The fit of our model was good with Pseudo R^2 being 0.40 (Table 4). Differences between participants and refusers were primarily explained by prosocial intrapersonal characteristics, while demographic characteristics provided only limited explanation. Notably, in our model, having a (registered) partner and a low educational level were relevant but slightly weaker predictors among the demographic variables. Low educational level significantly increased the odds to be participant in Lifelines in comparison with moderate or high educational levels, while a high level compared to moderate and low level of education had a smaller yet significant effect. In contrast, residence did not any longer contribute to differences between participants and refusers, once other predictors were taken into account (OR = 0.76, $p = 0.10$). A higher self-reported health significantly predicted participation in our model (OR = 1.28, $p = 0.02$).

Healthspheric values were the strongest predictors of participation in the simultaneous analyses of demographic and psychological predictors. This finding is reflected in the significant predictors of blood donor status ($OR = 2.31, p < 0.01$) and organ donor status ($1.94, p < 0.01$). The lower scores in biospheric values of participants compared to refusers were robust in our multivariate analyses (3.76 vs. $3.83, p < 0.01$). Table 4 shows that associations of SVO score ($OR = 1.00, p = 0.37$) and frequency of charitable donations ($OR = 1.10, p = 0.20$) with participation in the univariate analyses were not robust in the multivariate model. Although societal trust as well as trust in data management and handling by the government independently contributed to the prediction of participation, the level of trust in data management and handling by commercial researchers was a key predictor of participation ($OR = 0.57, p = 0.01$). There is a clear lack of trust among participants, while refusers seem to have less issues in trusting commercial researchers and personal data.

Table 4. Results of the logistic regression of participants and refusers of the Lifelines biobank

Variable	Step	OR	95% CI OR	P	Pseudo R ²
					[R ² Cox & Snell; R ² Nagelkerke]
Age	1	1.01	[1.00, 1.02]	0.09	[0.03;0.06]
Gender	1	0.9	[0.68, 1.19]	0.45	
Marital status	1	1.68	[1.24, 2.28]	< 0.01	
Educational level High versus Other	1	0.58	[0.42, 0.79]	< 0.01	
Educational level Low versus Other	1	1.77	[1.16, 2.71]	< 0.01	
Paid job (yes = 1, no = 0)	1	1.32	[0.94, 1.85]	0.11	
Religious (yes = 1, no = 0)	1	1.06	[0.79, 1.41]	0.72	
Residence (rural = 1 urban = 2)	1	0.81	[0.58, 1.12]	0.2	
Self-reported health	1	1.28	[1.04,1.57]	0.02	

Frequency of charitable donations	2	1.1	[0.95, 1.27]	0.2	[0.04;0.07]
Blood donor	3	2.31	[1.35, 3.94]	< 0.01	[0.04;0.08]
Organ donor	4	1.94	[1.45, 2.58]	< 0.01	[0.05; 0.10]
Social Value Orientation score	5	1	[1.00, 1.01]	0.37	[0.05;0.10]
Healthsperic values	6a	-	-	0.02	[0.11;0.20]
Healthsperic values ²	6b			< 0.01	[0.13;0.25]
Healthsperic values ³	6c			< 0.01	[0.14;0.27]
Hedonic values	7a	-	-	< 0.01	[0.15;0.28]
Hedonic values ²	7b			< 0.01	[0.17;0.33]
Biospheric values	8a	-	-	< 0.01	[0.18;0.35]
Biospheric values ²	8b			< 0.01	[0.19;0.36]
Societal trust	9	1.38	[1.08, 1.77]	< 0.01	[0.19;0.37]
TDM&H/government	10a	-	-	< 0.01	[0.20;0.38]
TDM&H/government ²	10b			< 0.01	[0.20;0.38]
TDM&H/commercial enterprises	11	0.75	[0.61, 0.93]	< 0.01	[0.20;0.38]
TDM&H/commercial researchers	12	0.57	[0.45, 0.73]	< 0.01	[0.21;0.40]

Note: OR = Odds Ratio (not reported if transformed), CI = Confidence interval, Test statistics multivariate model final step: $\chi^2(26) = 615.414$, $p < 0.01$

Discussion

In this study, we examined demographic and psychological predictors of (non-)participation within a comparative analysis of biobank's participants and refusers. The results indicated that prosocial intrapersonal variables were stronger predictors of participation in biobanking than demographic variables. Specifically, health-related values and trust in management and handling of data by researchers appeared as key factors in distinguishing participants from refusers.

Our results add to previous findings on the importance of prosocial values and societal trust in participation in biobanking [3, 4, 6, 13, 18, 41]. The appraisal of the context of biobanking, that is reflected in prosocial values and trust in society or government, was a key factor for participation. First, we found opposite associations of healthspheric and biospheric values with participation. This indicates that refusers prioritize health or healthcare differently. The differences in donor status for blood or organs support this explanation, since the odds ratio to be blood or organ donor was smaller for refusers. Even though studies showed that motives for research participation are prosocial [18, 42], we did not find a stronger general "prosociality" score of participants compared to refusers when taking into account all other factors. Similarly, a recent psychological study with college majors found associations of organ donation with a medicine major, but only positive trends for prosocial SVO score [43].

Second, the opposite associations of trust in data management and handling of government and of commercial researchers with participation suggest that trust could simultaneously work as facilitator and barrier for participation in biobanking depending on the context. Refusers in our study had less societal trust yet more trust in commercial researchers compared to participants. This lower variance in refusers' trust in public vs. private context might indicate a more sceptic yet homogenic perception regardless of context. Two other studies found similar variances of trust in these contexts among non-participants [5, 30]. Moreover, our results extend findings of a recent international comparison of trust and intention to participate in genomic research and big data initiatives, since we show how participants and refusers perceive systems of authority differently, especially in research [5, 6]. Our findings argue for a context-specific understanding of trust [30, 44]. Similar to previous findings, trust in biobanks depended on its context, e.g., biobanks' funding stream or accessibility for third-party researchers [3, 4, 6, 45].

In our analyses, only educational level and being with partner were relevant demographic predictors of participation when taking into account prosocial in-

trapersonal characteristics. This supports previous findings on associations of participation with higher educational level and having a marital status [4–6]. First, a recent multinational study about genetic data sharing showed a similar effect for marital status on participation intention [6]. Recruitment aiming to include heritability might be more prone for inclusion of citizens with partner. Recruitment strategies need to address this bias separately, given its robust effect in our multivariate model.

Second, we found that a higher education was associated with a higher probability of refusing to participate, regardless of levels of trust. This finding contradicts the literature showing that higher educational level is associated with participation [5, 6], yet confirms a recent study showing a highest willingness to participate in biobank research for middle educational level [22]. Possibly, our finding can be explained as a methodological issue, since the refusers sample underrepresented low educational level, while the participants sample represented this group more accurately. It is, however, also possible that psychological or social mechanisms explain this finding; samples with a mandatory sampling strategy provide better representativeness than voluntary samples [46]. The peer/family pressure for three generations of one family participating in Lifelines might have given the Lifelines' sample a mandatory character, which was not the case for the refusers panel. Another explanation might be a variance in motives for (non-)participation among higher and lower educated people, or a relevant variance in their social networks, including other people participating. Future research should investigate how educational level and marital status independently might affect participation and refusal of participation via these or other psychological or social factors.

Implications

Our findings suggest that biobanks seeking to recruit participants should consider multiple prosocial intrapersonal characteristics within their recruitment strategies. Our findings indicate that healthspheric, biospheric values and trust in government are strongly associated with participation in large-scale biomedical research. Therefore, emphasizing contextual prosocial benefits of participation and increasing the trustworthiness of biobanks as being independent from the government may facilitate the recruitment of new participants. To retain participants, it may be wise to be cautious regarding collaborations with commercial enterprises because of participants' distrust of commercial researchers. The lack of trust in commercial enterprises is a general concern for recruitment and retention of participants. Our results highlight the importance of establishing governance and accountability structures that enhance trust

in the biobanking context [45]. Solutions might be found in introducing trustworthy researchers as gatekeepers to data, in particular hospital researchers [6]. Although the role of societal trust has been acknowledged, research practices that strengthen trust in different private and public contexts merit greater attention. This consideration is particularly pressing in the field of personalized medicine, in which the collaborative management and handling of data, entailing diverse contexts, have sparked serious concerns over data use [18, 29].

Limitations

Although our study has provided insights into differences between participants and refusers of medical research, it has some limitations. Sampling bias may have occurred, as individuals who refused to participate in the biobank were nevertheless willing to complete an online questionnaire on this topic. Consequently, they may not be representative of all individuals unwilling to participate in biobanks or data repositories. Our sample of refusers was experienced and aware of commercial research or the procedures for research in general, which may explain their higher levels of trust in commercial researchers [4]. Nevertheless, our findings provide valuable inputs for recruitment strategies and effective methods for data repository.

Conclusion

In conclusion, we have shown that participants and refusers of biobanking can be distinguished according to their prosocial intrapersonal characteristics. Our findings may contribute to improving recruitment strategies by incorporating relevant values and/or highlighting prosocial benefits. Moreover, they underline the need to address trust issues in collaborations between data repositories and commercial companies. Future research should explore how prosocial intrapersonal characteristics drive participation and withdrawal decisions.

References

1. Rivera SM, Brothers KB, Cadigan RJ, Harrell HL, Rothstein MA, Sharp RR, et al. Modernizing research regulations is not enough: it's time to think outside the regulatory box. *Am J Bioeth.* 2017;17:1–3. <https://doi.org/10.180/15265161.2017.1328899>.
2. Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L. The ethics of algorithms: mapping the debate. *Big Data Soc* 2016;2016:68. <https://doi.org/10.1177/2053951716679679>
3. Gaskell G, Gottweis H, Starkbaum J, Gerber MM, Broerse J, Gottweis U, et al. Publics and biobanks: Pan-European diversity and the challenge of responsible innovation. *Eur J Hum Genet.* 2013;21:14–20. <https://doi.org/10.1038/ejhg.2012.104>.
4. Critchley C, Nicol D, Otlowski M. The impact of commercialisation and genetic data sharing arrangements on public trust and the intention to participate in biobank research. *Public Health Genomics.* 2015;18:160–72. <https://doi.org/10.1159/000375441>.
5. Middleton A, Milne R, Thorogood A, Kleiderman E, Niemiec E, Prainsack B, et al. Attitudes of publics who are unwilling to donate DNA data for research. *Eur J Med Genet.* 2019;62:316–23. <https://doi.org/10.1016/j.ejmg.2018.11.014>.
6. Milne R, Morley KI, Howard H, Niemiec E, Nicol D, Critchley C, et al. Trust in genomic data sharing among members of the general public in the UK, USA, Canada and Australia. *Hum Genet.* 2019;138:1237–46. <https://doi.org/10.1007/s00439-019-02062-0>.
7. Thompson A. Thinking big: large-scale collaborative research in observational epidemiology. *Eur J Epidemiol.* 2009;24:727–31. <https://doi.org/10.1007/s10654-009-9412-1>.
8. Kettis-Lindblad Å, Ring L, Viberth E, Hansson MG. Genetic research and donation of tissue samples to biobanks. What do potential sample donors in the Swedish general public think? *Eur J Public Health.* 2006;16:433–40. <https://doi.org/10.1093/eurpub/ckl198>.
9. Critchley C, Nicol D, Otlowski M, Stranger M. Predicting intention to biobank: a national survey. *Eur J Public Health.* 2012;22:139–44. <https://doi.org/10.1093/eurpub/ckq136>.
10. Ahram M, Othman A, Shahrouri M, Mustafa E. Factors influencing public participation in biobanking. *Eur J Hum Genet.* 2014;22:445–51. <https://doi.org/10.1038/ejhg.2013.174>.

11. Kaufman D, Murphy J, Scott J, Hudson K. Subjects matter: a survey of public opinions about a large genetic cohort study. *Genet Med*. 2008;10:831–9. <https://doi.org/10.1097/GIM.0b013e31818bb3ab>.
12. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the LifeLines cohort study. *PLoS ONE*. 2015;10:1–12. <https://doi.org/10.1371/journal.pone.0137203>.
13. Sanderson SC, Brothers KB, Mercaldo ND, Clayton EW, Antommaria AHM, Aufox SA, et al. Public attitudes toward consent and data sharing in biobank research: a large multi-site experimental survey in the US. *Am J Hum Genet*. 2017;100:414–27. <https://doi.org/10.1016/j.ajhg.2017.01.021>.
14. Merdad L, Aldakhil L, Gadi R, Assidi M, Saddick SY, Abuzenadah A, et al. Assessment of knowledge about biobanking among healthcare students and their willingness to donate biospecimens. *BMC Med Ethics*. 2017;18:32. <https://doi.org/10.1186/s12910-017-0195-8>.
15. Rahm AK, Wrenn M, Carroll NM, Feigelson HS. Biobanking for research: a survey of patient population attitudes and understanding. *J Community Genet*. 2013;4:445–50. <https://doi.org/10.1007/s12687-013-0146-0>.
16. Ridgeway JL, Han LC, Olson JE, Lackore KA, Koenig BA, Beebe TJ, et al. Potential bias in the bank: what distinguishes refusers, nonresponders and participants in a clinic-based biobank? *Public Health Genomics*. 2013;16:118–26. <https://doi.org/10.1159/000349924>.
17. Goddard KABB, Smith KS, Chen C, McMullen C, Johnson C. Biobank recruitment: motivations for nonparticipation. *Biopreserv Biobank*. 2009;7:119–21. <https://doi.org/10.1089/bio.2009.0006>.
18. Richter G, Borzikowsky C, Lieb W, Schreiber S, Krawczak M, Buyx A. Patient views on research use of clinical data without consent: legal, but also acceptable? *Eur J Hum Genet*. 2019;27:841–7. <https://doi.org/10.1038/s41431-019-0340-6>.
19. Luhmann N. Trust and power. John Wiley & Sons Inc.; 1979. p. 228.
20. Rotter JB. Generalized expectancies for interpersonal trust. *Am Psychol*. 1976;26:443–52.

21. Porteri C, Pasqualetti P, Togni E, Parker M. Public's attitudes on participation in a biobank for research: an Italian survey. *BMC Med Ethics*. 2014;15:81. <https://doi.org/10.1186/1472-6939-15-81>.
22. Bossert S, Kahraß H, Strech D. The public's awareness of and attitude toward research biobanks A regional German survey. *Front Genet*. 2018;9:1–11. <https://doi.org/10.3389/fgene.2018.00190>.
23. Critchley C, Nicol D, McWhirter R. Identifying public expectations of genetic biobanks. *Public Underst Sci*. 2017;26:671–87. <https://doi.org/10.1177/0963662515623925>.
24. Hyde MK, Knowles SR. What predicts Australian university students' intentions to volunteer their time for community service? *Aust J Psychol*. 2013;65:135–45. <https://doi.org/10.1111/ajpy.12014>.
25. van der Werff E, Steg L. The psychology of participation and interest in smart energy systems: comparing the value-belief-norm theory and the value-identity-personal norm model. *Energy Res Soc Sci*. 2016;22:107–14. <https://doi.org/10.1016/j.erss.2016.08.022>.
26. Bouman T, Steg L, Kiers HAL. Measuring values in environmental research: a test of an environmental portrait value questionnaire. *Front Psychol* 2018;9:1–15. <https://doi.org/10.3389/fpsyg.2018.00564>.
27. Steg L, Perlaviciute G, van der Werff E, Lurvink J. The significance of hedonic values for environmentally relevant attitudes, preferences, and actions. *Environ Behav*. 2014;46:163–92. <https://doi.org/10.1177/0013916512454730>.
28. Evers A, Gesthuizen M. The impact of generalized and institutional trust on donating to activist, leisure, and interest organizations: Individual and contextual effects. *Int J Nonprofit Volunt Sect Mark*. 2011;16:381–92. <https://doi.org/10.1002/nvsm.434>.
29. Raivola V, Snell K, Helén I, Partanen J. Attitudes of blood donors to their sample and data donation for biobanking. *Eur J Hum Genet*. 2019;27:1659–67. <https://doi.org/10.1038/s41431-019-0434-1>.
30. Broekstra R, Aris-Meijer J, Maeckelberge E, Otten S, Stolk R. Trust in

- centralised large-scale data repository: a qualitative analysis. *J Empir Res Hum Res Ethics*. 2019;1–14. <https://doi.org/10.1177/1556264619888365>.
31. Mújdricza F. The roots of trust. *Eur J Ment Heal*. 2019;14:109–42. <https://doi.org/10.5708/EJMH.14.2019.1.6>.
 32. Scholtens S, Smidt N, Swertz MA, Bakker SJLL, Dotinga A, Vonk JM, et al. Cohort profile: LifeLines, a three-generation cohort study and biobank. *Int J Epidemiol*. 2015;44: 1172–80.
 33. Messick DM, McClin, McClintock CG. Motivational bases of choice in experimental games. *J Exp Soc Psychol*. 1968;4:1–25. [https://doi.org/10.1016/0022-1031\(68\)90046-2](https://doi.org/10.1016/0022-1031(68)90046-2).
 34. Murphy RO, Ackermann KA, Handgraaf MJJ. Measuring social value orientation. *Judgm Decis Mak*. 2011;6:771–81. <https://doi.org/10.2139/ssrn.1804189>.
 35. Schwartz SH, Cieciuch J, Vecchione M, Davidov E, Fischer R, Beierlein C, et al. Refining the theory of basic individual values. *J Pers Soc Psychol*. 2012;103:663–88. <http://doi.apa.org/getdoi.cfm?doi=10.1037/a0029393>.
 36. Mayer RC, Davis JH, Schoorman FD. An integrative model of organizational trust. *Acad Manag Rev*. 1995;20:709–34. <http://www.jstor.org/stable/258792>.
 37. IBM Corp. IBM SPSS statistics for Windows, version 25.0. Armonk, NY: IBM Corp.; 2017.
 38. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate data analysis. 7th ed. Pearson Prentice Hall; 2014. p. 734.
 39. Box GEP, Tidwell PW. Transformation of the independent variables. *Technometrics*. 1962;4:531–50. <https://www.tandfonline.com/doi/abs/10.1080/00401706.1962.10490038>.
 40. Peduz P, Concato J, Kemper E, Holford TR, Feinstein AR. Study of the number of events per variable. *Logist Regres Anal*. 1996;49:1373–9.
 41. Gayet-Ageron A, Rudaz S, Perneger T. Biobank attributes associated with higher patient participation: a randomized study. *Eur J Hum Genet*.

- 2016;25:31–6. <https://doi.org/10.1038/ejhg.2016.132>.
42. Locock L, Boylan AMR. Biosamples as gifts? How participants in biobanking projects talk about donation. *Heal Expect*. 2016;19:805–16. <https://doi.org/10.1111/hex.12376>.
 43. van Andel CEE, Tybur JM, Van Lange PAM. Donor registration, college major, and prosociality: differences among students of economics, medicine and psychology. *Pers Individ Dif*. 2016;94:277–83. <https://doi.org/10.1016/j.paid.2016.01.037>.
 44. Sheikh Z, Hoeyer K. “That is why I have trust”: unpacking what ‘trust’ means to participants in international genetic research in Pakistan and Denmark. *Med Heal Care Philos*. 2018;21:169–79. <https://doi.org/10.1007/s11019-017-9795-9>.
 45. Nicol D, Critchley C, McWhirter R, Whitton T. Understanding public reactions to commercialization of biobanks and use of biobank resources. *Soc Sci Med*. 2016;162:79–87. <https://doi.org/10.1016/j.socscimed.2016.06.028>.
 46. Cheung KL, Ten Klooster PM, Smit C, De Vries H, Pieterse ME. The impact of non-response bias due to sampling in public health studies: a comparison of voluntary versus mandatory recruitment in a Dutch national survey on adolescent health. *BMC Public Health*. 2017;17:1–10. <https://doi.org/10.1186/s12889-017-4189-8>.

Chapter 5

Motives for withdrawal of participation in biobanking and participants' willingness to allow linkages of their data

Abstract

Data repositories, like research biobanks, seek to optimize the number of responding participants while simultaneously attempting to increase the amount of data donated per participant. Such efforts aim to increase the repository's value for its uses in medical research to contribute to improve health care, especially when data linkage is permitted by participants. We investigated individuals' motives for participating in such projects and potential reasons for their withdrawal from participation in a population-based biobank. Additionally, we analysed how these motives were related to various characteristics of the participants and their willingness to permit data linkage to their personal data for research. These questions were explored using a sample of participants in the Dutch Lifelines biobank ($n = 2,615$). Our results indicated that motives for participation and withdrawal were premised on benefits or harm to society and to the individuals themselves. Although general values and trust both played key roles in participation, potential withdrawal, and willingness to permit data linkage, they were differentially associated with motives for participation and withdrawal. These findings support and nuance previous findings by highlighting the distinctiveness and complexity of decision making regarding participation in or withdrawal from data donation. We suggest some new directions for improving recruitment, retention and safeguarding strategies in biobanking. In addition, our data provide initial evidence regarding how factors may relate with the probability that individuals will agree to data linkages, when controlling for their unique effects. Future research should further investigate how perceptions of harm and benefits may influence decision making on withdrawal of participation.

Keywords: biobanks; participation; withdrawal; data linkage; commercialization

This chapter is currently accepted for publication by the European Journal of Human Genetics on behalf of Broekstra, R., Aris, J., Maeckelberghe, E.L.M, Stolk, R.P., Otten S.

Introduction

Continuous and full participation of members of the public in data repositories in the medical field is essential for effective scientific research. In addition, it creates the opportunity to improve health care for individuals by data enhancement [1–3]. To understand complex relationships in retaining health or developing disease, a large number of participants contributing large amounts of data is needed to get sufficient statistical power [3,4]. Analyses on a large scale, centralized data repository enables researchers to understand specific characteristics and behaviour on an individual level in an unprecedented way, which can impact both research and clinical practice [4]. Additionally, advances in information technology create the possibility to analyse vast amounts of genetic (genomic) data, and the potential to understand more about risk factors for developing diseases or the course of the disease. As a result of these new possibilities for analysing and knowledge building, the quest to advance the collection and use of data from both patients and healthy volunteers has become urgent, further increasing the efforts on exponential growth of data repository and international data-sharing or linkage projects [5,6].

Large-scale, centralized data repositories can be large prospective population-based cohort studies and biomedical biobanks. These biomedical cohort studies and biobanks (further referred to as biobanks) store large quantities of biological specimens as well as data extracted from questionnaires and measurements as a resource for research, while being managed by professional standards [7,8]. Biobanks and similar data repositories seek to optimize the number of responding participants increasing the data repositories' value for health care and research, while simultaneously attempting to increase the amount of data donated per participant. Yet recruiting participants in biobanking in Europe is a challenge even in 'willing' populations, for example among Finnish people [9]. Achieving high participation rates will most likely become an even larger challenge, considering the development of participation rates in population-based survey studies in the past 35 years [10,11]. In light of these challenges for recruiting new participants, retainment of current participants should be a growing concern for those responsible for data repositories and for those using them.

The establishment of mutually enhancing data linkages between existing data repositories might be an important solution for improving the collection and use of data. Data linkage could provide multiple opportunities, such as providing unique insights in existing data without significant efforts of participants [5], or enabling individually tailored insights from data that might be communicated in return for participation to benefit a participants' health directly [3].

As a result, data linkage might influence current rates of participation and data density, especially when individually tailored results are returned or accessible as return of an individual benefit.

Yet data linkage is challenging when it comes to protection of personal sensitive data, and promising potential individual benefits [1,5]. The clinical utility of individually tailored insights from research is not straight forward and an ongoing topic of debate, for example the clinical utility of individuals' polygenic risk scores for various illnesses [12] or of data from wearables [13]. Research has shown that motives for non-participation and (potential) withdrawal are linked to issues regarding commercialisation, confidentiality and privacy of data [14]. Simultaneous, biobank's operating procedures, such as restricted data access policies, or its decision to return findings from research are associated with higher participation levels [15,16].

Achieving optimized rates of participation therefore requires a sound understanding of citizens' decisions to participate or not-participate in biobanks. Especially understanding decisions of withdrawal might help to understand differences between participant and (potential) non-participant. Subsequently, it provides valuable insights in the viability of data sharing as potential solution for optimized rates. Yet information about rates of withdrawal in biobanking is scarce, or not publicly available, which limits current understanding of mechanisms of withdrawal. Although one Swedish study investigating withdrawal rates in biobanking during 2005 and 2006 indicated low withdrawal rates at that time [17], the potential of data and data linkage have rapidly developed, and concerns about data linkage in the last decade became realities [18].

So far, studies have pointed to the key roles and interplay of individuals' characteristics in their participation and non-participation in biobanking. For example, a variety of characteristics on an individual level have been found to be positively associated with participation in biobanks, such as a higher educational attainment [19–21], positive attitudes towards research and society [6,20,22], a general concern about others [23], and high levels of societal or social trust [6,21]. In addition, several studies suggest that prosocial attitudes and trust in others are especially important factors that independently relate to positive decisions regarding participation [24,25].

Some of these characteristics may also be positively related to withdrawal of participation, depending on the specific context of the biobank, e.g. studies in Europe, United States, and Australia showed that concerns about privacy and confidentiality were robustly related to lower willingness to participation [6,26,27]. Specifically these concerns could be triggered in case of data linkage or access, especially when commercial enterprises are involved [28,29]. An Aus-

tralian study showed that higher educational attainment was associated with a stronger reduction of trust in a public biobank that allows access to third parties [30]. Other studies found that those concerns about privacy or confidentiality were associated with anxiety for commodification of contributed samples or data [31,32]. Furthermore, technologies are perceived differently depending on health expectancies, gender, or social trust [33]. For example, men perceive fewer risk and more benefits in gene technology than women [34]. Levels of trust in data linkage or commercial enterprises can also inhibit or decrease perceived risk in participation, which might be decisive for continuation or withdrawal [21,30].

Hence, the contextual and procedural characteristics of biobanks can affect participation via withdrawal, depending on the characteristics of (potential) participants. Changes in the data after withdrawal can lead to bias in the data and the analysis of the data from data repositories, especially in biobanks representing populations of both participants and non-participants, which could be reduced by refinement of the data [25,35]. While previous studies have broadly explored the characteristics and motives of participants and non-participants in biobanks as well as in data sharing [6,20,21,26,36,37], further research on the associations between individual participants' characteristics and their motives for participation and withdrawal would be beneficial [35]. In particular, research on the associations between participants' motives of withdrawal and their psychological characteristics could yield insights into how contextual biobanking elements could impact on individuals' continued participation or their withdrawal [24,25].

An investigation of the drivers of participation and withdrawal behaviour is therefore required to reduce potential biases. This study aimed to gain insight in the possibilities to classify individuals' who participate in population-based biobanks according to various demographic and psychological key characteristics, and find distinguishing traits between those participants who are more likely to withdraw their participation and those who are not. In addition, we tried to identify factors that will affect the likelihood that participants would accept the linkage of their personal data for research purposes. The study could then yield insights for optimizing recruitment and retainment strategies associated with biobanking or similar data repository initiatives, while preventing withdrawal of participation.

Materials and Methods

Procedure and participants

We administered an online survey in August 2018 within a randomly selected sample of 2,615 respondents, who were among the 167,000 participants of Lifelines biobank. Lifelines is a population-based biobank and large, multigenerational prospective cohort study in the Northern part of the Netherlands [7]. All Lifelines participants consented between 2006 and 2013 to participate in the biobank, allowing it to examine 10 percent of the population of Northern part of the Netherlands. It does not offer incentives for research participation. Lifelines applies a broad range of investigative procedures for assessing biomedical, sociodemographic, behavioural, physical, and psychological factors, with a particular focus on multi-morbidity and complex genetics.

At the time of data collection for the current study, our respondents were registered as active participant of Lifelines, which implies they had not withdrawn their participation at the time of filling in our survey and were actively participating in the biobank. The response rate for the invitation of this additional study was 22.2%. We stratified our sample to improve its representativeness compared to the representativeness of the Lifelines cohort for the Dutch population [7,38]. We found that our study sample, stratified by sex (male = 50.5%) and age ($M = 56$, $SD = 15.88$), nevertheless has more individuals with a high educational level, registered partner, and good self-reported health compared to the Lifelines population [38].

Measures

Demographic characteristics and self-reported health

We measured demographic characteristics that were previously associated with participation or willingness to participate in a biobank, such as marital status [21], education level [19,24], religion [16,36], residential area [38], and self-reported general health [19,39]. Work status was included as an additional indicator of socio-economic status.

Prosocial Intrapersonal Characteristics

Prosocial Orientation.

We applied two measures to determine respondents' general orientations towards other individuals within their behaviour, which we refer to as prosocial orientation. The first was their actual prosocial behaviour reflected in their organ and blood donor status and the frequency of charitable donations. The second

was the degree to which respondents cared about other individuals' outcomes in choices concerning resource allocation, which were measured using six items in the social value orientation (SVO) scale. This scale is designed to measure the magnitude of individuals' general concern for others [40]. Replicating previous research, our results indicated excellent reliability with Cronbach's alpha > 0.90 [41].

Values.

We used 11 items based on the Theory of Basic Values from Schwartz to assess motivations of and beyond self-interest as guiding principles in individuals' life [42]. For example, about the relevance of individual pleasure in life as guiding principle or the relevance of health and healthcare. Although our primary focus was on motivations of and beyond self-interest in the healthcare context, we investigated these motivations relating to a wider context, such as natural environmental values [43]. As such, values can provide more information about both magnitude and direction of individuals' motives of self-interest and beyond. We applied an adapted version of the Environmental Personal Values Questionnaire (E-PVQ) scale [44]. We used two different subscales from the E-PVQ to measure hedonic values associated with (individual pleasure) (e.g., 'it is important to have fun') and biospheric values relating to the natural environment (e.g., 'it is important to protect the natural environment'). In addition, we used a tailored subscale on healthspheric values relating to health and healthcare concerns (e.g., 'it is important to live a fit life'). This subscale was designed in collaboration with the developers of the E-PVQ. The scales' reliability was good, with Cronbach's alpha > 0.80 for all of the subscales, thus confirming reliability [44].

Societal trust.

We identified trust as a key factor of participation and nonparticipation in biobanking. This trust can relate to the general trust in society and a domain specific trust in research, especially in case of data linkage in research. Additionally, a previous study distinguished levels of trust in organizations from levels of trust in employees of these organisations [45]. We therefore investigated both the level of individuals' trust in society, and the level of trust in research organisations, in particular their data management and handling (hereinafter referred to as 'DM&H'). The latter should automatically address issues about privacy in data management and handling. First, societal trust was measured using the trust-based framework proposed by Mayer, Davis, and Schoorman [46]. We used six items reflecting trust in the government and in other citizens that were measured with a 5-point Likert scale. Examples included 'the government acts with good intentions'; 'people in the Netherlands are willing to help each oth-

er'; and 'people are trustworthy'. These items focused on aspects of societal or sometimes called social trust, namely trustworthiness, good intentions, and competence. The combined scale with all six items demonstrated good reliability (Cronbachs' $\alpha = 0.86$).

Second, we used five items focusing on trust in the DM&H of different types of research organizations (hospitals, governmental institutes, universities, large-sized commercial enterprises, and small and medium-sized enterprises (SMEs)). An illustrative item was 'I believe that the hospital correctly, adequately, and fairly stores and treats my personal data'. Besides measuring trust in organizations as entities, we also measured trust in specific research employees using four items (hospital researchers, university researchers, commercial market researchers, and commercial polling researcher). An illustrative item was 'I trust employees of commercial polling enterprises to conduct research using a correct, fair and careful approach'.

We conducted a principal component analysis with an oblimin rotation given several high intercorrelations between several research items ($r > 0.50$). The results of the analysis revealed that the items could be reduced to three factors explaining 71.90% of the total variance. The first factor, comprising four items, was trust in DM&H of the government (organizations and their employees) and explained 45.74% of the total variance ($\alpha = 0.83$). The second factor, comprising two items, was trust in the DM&H of commercial enterprises and explained 14.24% of the total variance ($r = 0.65$). The third factor, trust in DM&H of commercial researchers, comprised two items and explained 11.92% of the total variance ($r = 0.47$). Correlations between the components were low to moderate ($r < 0.40$).

Motives for participation and potential withdrawal.

Although we included values to assess motivations of and beyond self-interest, a better understanding of motives would require including the biobanking context specific motives. We measured therefore a variety of potentially relevant motives for participation and withdrawal in biobanking derived from an earlier explorative qualitative study on this topic [45]. These motives included topics of potential benefits and potential harms of participation in biobanking. We performed a factor analysis of nine motives for participation (e.g., 'I contribute to health care by participating') and seven motives for potential withdrawal (e.g., 'I had a negative experience during a visit'). We conducted two distinct principal axis factor analyses with oblimin rotation, which revealed that motives for participation and withdrawal could be clearly differentiated according to two factors. Yet, three of our initial items about potential motives were excluded for

further analyses due to their low communalities (< 0.30). The appropriateness of these factors relating to remaining motives for participation and withdrawal were confirmed by the results of Kaiser-Meyer-Olkin tests (0.795 and 0.785) and Bartlett's tests of sphericity ($\chi^2(120) = 13,557.87$, $p < 0.01$; $\chi^2(21) = 5,025.84$, $p < 0.01$), respectively. The two factors relating to participation, that is, 'societal benefits' and 'individual benefits', respectively explained 45.19% and 17.91% of the total variance. The two factors relating to potential withdrawal, that is, 'societal harm' and 'individual harm', respectively explained 37.06 % and 8.89 % of the total variance.

Willingness of participants to allow personal data linkages.

We measured the participants' willingness to accept data linkage of five types of potentially relevant personal data (e.g., medical, financial, and sensor data) for large-scale scientific research using a 5-point Likert scale. Data linkages were first introduced to the participants with several illustrations of their application, for example in electronic health records or annual tax income report. Next, participants were asked the following question: 'How likely would it be for you to share the following data for large-scale scientific research?' (1 = 'very unlikely', 5 = 'very likely'). The total score ranged from 5 to 25 and served as an indicator of participants' willingness to agree to the linking and sharing of their personal data to a biobank.

Data analysis

We calculated descriptive statistics for demographic and psychological variables (motives and prosocial orientations). We applied Chi-square tests, (M) ANOVA tests, or independent t-tests to calculate potential differences in means for motives to participate or withdraw among different demographic groups. Further, we performed bivariate correlational and linear multiple regression analyses, applying a manual entry strategy to investigate associations between motives and individual characteristics. We entered the variables forward block wise following the order of our description of measurements. The statistical power of this sample was high (1.00), and an a-priori conservative power analysis with Gpower software showed a minimum of $n = 1,347$. All other analyses were conducted using the SPSS software, version 25.0 [47].

Results

Before proceeding to our analyses, the quality of the data was assessed by explorative descriptive analyses of missing data. The assessment showed less than five percent missing data per variable. Even though there was some miss-

ing data, there was no clear pattern in its missingness. Hence, we believed this data was missing at random. We proceeded with descriptive and explorative analyses with univariate comparisons. Subsequently, we conducted our main analyses of predicting motives of withdrawal and intention to allow participants' data to be linked.

Demographic and prosocial characteristics

Table 1 depicts the participants' demographic and prosocial characteristics. Of the 2,615 respondents, 79.5% had registered partnerships, such as marriages, and 55.8% had paid jobs. The participants reported being generally in good health, with a mean score of 3.97 on a 5-point scale. A total of 41% of the respondents were highly educated (n = 1083), and 37% were religious (n = 972). The majority of the participants (77 %) lived in rural areas (n = 2019). Further, 12% of the respondents were blood donors and 56% registered as organ donors. The mean value for contributions to charities was moderate at 2.68 on a 5-point scale. The SVO score for participants' concern about others was 32.25, indicating a strong prosocial orientation [41]. In addition, Table 1 shows most participants scored higher for values relating to health, a healthy lifestyle, and their own pleasure than for environmental values. They reported high levels of trust in the government and in other citizens (M = 3.25, SD = 0.63), and especially in the government's DM&H (M = 3.55, SD = 0.81). However, they had lower levels of trust in the DM&H practices of commercial researchers (M = 3.32, SD = 0.80) and commercial enterprises (M = 3.14, SD = 0.82).

Table 1. Demographic characteristics and the self-reported health of respondents being participants in a Dutch population-based biobank

	N / M	% / SD
Age (categorical)		
< 30 years	231	8.80%
31 – 40 years	252	9.70%
41 – 50 years	357	13.70%
51 – 60 years	580	22.20%
> 60 years	1192	45.60%
Marital status		
Registered partner	2,078	79.50%

No registered partner	536	20.50%
Educational level		
High	1,083	41.40%
Moderate	938	35.90%
Low	559	21.70%
Paid job	1,460	55.80%
Religious	972	37.20%
Area of residence		
Rural	2,019	77.20%
Urban	537	20.50%
Self-reported health (1= very poor, 5 = very good)	3.97	0.71
Frequency of donations to charities	2.68	1.08
Blood donor	306	11.70%
Organ donor	1,460	55.80%
Social Value Orientation score	32.25	14.19
Values		
Healthspheric	3.99	0.66
Hedonic	4.25	0.65
Biospheric	3.76	0.75
Societal trust	3.25	0.63
TDM&H/ government	3.55	0.81
TDM&H/ commercial enterprises	3.14	0.82
TDM&H/ commercial researchers	3.32	0.8

Note: The values in the table are either means (M) with standard deviations (SD) or frequency numbers (N) with percentages (%) relating to the total sample. TDM&H: TDM = trust in data management; and H=handling. All variables' missing data < 5%.

Motives for participation and withdrawal

Table 2 provides the results of our analyses of motives for participation and potential withdrawal, respectively. Overall, respondents scored higher on societal than for individual motives relating to both participation (4.06 vs 2.92) and

withdrawal (3.94 vs 3.49). Table 3 shows the Pearson’s correlations between motives for participation and (potential) withdrawal and intrapersonal characteristics. We found modest associations of motives ($r < 0.30$), and especially participation motivated by societal benefits, with the various indicators of individual prosocial values and trust. All of the trust indicators were positively related to participation motivated by societal as well as individual benefits.

Table 2. Motives for participation in and potential withdrawal from a Dutch population-based biobank

	M	SD	Loading
<i>Motives of Participation – Societal benefits</i>	4.06	0.5	($\alpha = 0.83$)
I contribute to the health of future generations by participating	4.23	0.59	0.74
I contribute to health care by participating	4.14	0.55	0.91
I contribute to science by participating	4.1	0.57	0.84
I contribute to society by participating	3.77	0.74	0.55
<i>Motives of Participation - Individual benefits</i>	2.92	0.8	($\alpha = 0.84$)
I learn about myself by participating	3.86	0.75	-
I perceive participation as moral duty	3.22	0.97	0.47
I have fun by participating	2.89	0.95	0.74
People I know participate	2.71	1.21	-
I feel proud by participating	2.65	0.99	0.96
<i>Motives of Withdrawal - Societal harm</i>	3.94	0.84	($\alpha = 0.77$)
The research organisation becomes a for- profit organisation	3.98	0.89	0.87
The research organisation starts collaboration with commercial enterprises	3.91	0.93	0.83
The research does no longer contribute to health care	3.65	0.85	-
<i>Motives of Withdrawal - Individual harm</i>	3.94	0.84	($\alpha = 0.77$)
I have a negative experience during a visit	3.98	0.89	0.87
The research organisation gets negative news reports	3.91	0.93	0.83
De time investment becomes too high	3,48	0,91	0,59
The research aim changes	3,41	0,78	0,42

Note: The values in the table are means (M) with standard deviations (SD), with reliability statistic Cronbach’s alpha (α). Loadings of items with communalities < 0.30 are not reported (-), and were left out in further analyses.

Table 3. Correlations

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1. Motives of Participation - Societal benefits	1														
2. Motives of Participation - Individual benefits	0.38	1													
3. Motives of Potential Withdrawal - Societal harm	.04^s	-0.12	1												
4. Motives of Potential Withdrawal - Individual harm	-.05^s	-0.12	0.37	1											
5. Age	0.06	0.1	0.11	0.04	1										
6. Self-reported health	0.03	-.05^s	0.06	0.03	-0.07	1									
7. Frequency of charitable donations	0.11	-0.03	0.06	0.02	0.17	.05^s	1								
8. Social Value Orientation (SVO) score	0.08	-0.03	0.05	-0.01	0.01	0	0.13	1							
9. Healthspheric values	0.26	0.09	0.16	0.05	0.2	0.24	0.13	0.11	1						
10. Hedonic values	0.23	.05^s	0.11	0.05	-0.1	0.13	0.04	0.08	0.62	1					
11. Biospheric values	0.28	0.11	0.16	0.02	0.25	0.09	0.17	0.14	0.68	0.43	1				
12. Societal trust	0.15	0.05	0	0.03	0.02	0.22	0.16	0.12	0.18	0.11	0.15	1			
13. TDM&H/government	0.22	0.07	.04^s	-0.02	-.05^s	0.12	0.09	0.11	0.18	0.17	0.14	0.4	1		
14. TDM&H/commercial enterprises	0.06	0.2	-0.08	-0.06	0.06	0.03	0.07	-.05^s	0.06	0.02	0.02	0.24	0.32	1	
15. TDM&H/commercial researchers	0.14	0.13	-.04^s	-0.03	-.04^s	0.07	0.1	0.07	0.06	0.02	0.07	0.35	0.46	0.4	1

Note: **Bold** = significant with $p < .01$ **Bold^s** = significant with $p < .05$.

To obtain an overview of systematic associations of motives with demographic and psychological characteristics, we conducted an exploratory comparative analysis of means. Table 4 shows that motives relating to individual benefits as well as harm differed significantly yet modestly according to participants' educational levels. Motives to participate for individual benefit were stronger among participants with lower education levels ($M = 3.12$, $SD = 0.72$) than among moderately educated ($M = 2.98$, $SD = 0.81$) and highly educated participants ($M = 2.76$, $SD = 0.78$). However, the reverse was true for motives to withdraw participation relating to individual harm. These motives were less important for participants with lower education levels ($M = 3.45$, $SD = 0.65$) and moderately educated ($M = 3.49$, $SD = 0.62$) than they were for highly educated participants ($M = 3.54$, $SD = 0.54$).

In addition, we found that gender had a significant effect on the relevance of certain motives; whereas women scored lower than men for participation motivated by individual benefits, they scored higher for motives to withdraw participation based on societal and individual harm. Furthermore, blood and organ donors differed in their participation and potential withdrawal motives. Blood donor status simply predicted a stronger motivation to withdraw participation to avoid individual harm, whereas organ donor status was associated with higher scores for societal benefits and weaker scores for individual benefits as motives for participation. Finally, job status was significantly associated with all motives except for those relating to individual harm. The other demographic characteristics (i.e., marital status, religiosity, and urban dwelling) were not significantly associated with motives for participation and withdrawal.

Table 4. A comparison of means obtained for respondents' motives for participating in and potential withdrawing from a Dutch population-based biobank

	<u>Motives Participation</u>		<u>Withdrawal</u>	
	<u>Societal benefits</u>	<u>Individual benefits</u>	<u>Societal harm</u>	<u>Individual harm</u>
	p (Levine / F), BF / M (SD)	p (Levine / F), BF / M (SD)	p (Levine / F), BF / M (SD)	p (Levine / F), BF / M (SD)
Age	0.19 (1.50), n.a.	0.25 (1.33), n.a.	< 0.01 (3.94), n.a.	0.16 (1.60), n.a.
< 30 years	4.03 (0.48)	2.90 (0.79)	3.52 (0.68)	3.49 (0.59)
31 – 40 years	3.97 (0.59)	2.84 (0.80)	3.62 (0.72)	3.49 (0.62)
41 – 50 years	4.03 (0.53)	2.84 (0.85)	3.77 (0.65)	3.55 (0.58)
51 – 60 years	4.08 (0.50)	2.84 (0.84)	3.85 (0.72)	3.49 (0.65)
> 60 years	4.09 (0.47)	3.02 (0.75)	3.88 (0.75)	3.47 (0.66)
Gender	0.97 (0.00), 0.02	< 0.01 (6.86), 0.48	< 0.01 (19.16), 220.37	< 0.01 (26.88), 10 088.64
Male	4.06 (0.51)	2.97 (0.80)	3.88 (0.90)	3.43 (0.67)

Female	4.06 (0.49)	2.88 (0.80)	4.03 (0.77)	3.56 (0.60)
Marital status	0.11 (2.50), 0.06	0.35 (0.88), 0.02	0.35 (0.86), 0.02	0.17 (1.91), 0.04
Registered partner	4.07 (0.50)	2.93 (0.80)	3.95 (0.83)	3.50 (0.63)
No registered partner	4.03 (0.51)	2.89 (0.81)	3.91 (0.88)	3.46 (0.66)
Educational level	0.20 (1.63), 0.00	< 0.01 (44.54), 3.97E+15	0.05 (3.11), 0.01	< 0.01 (11.59), 39.92
High	4.08 (0.53)	2.75 (0.78)	3.99 (0.82)	3.56 (0.59)
Moderate	4.05 (0.47)	2.98 (0.77)	3.91 (0.84)	3.46 (0.66)
Low	4.04 (0.50)	3.12 (0.81)	3.90 (0.89)	3.41 (0.68)
Paid job	< 0.01 (9.81), 2.08	< 0.01 (19.83), 303.72	< 0.01 (7.03), 0.53	0.33 (0.94), 0.03
Yes	4.03 (0.52)	2.86 (0.81)	3.91 (0.84)	3.50 (0.60)
No	4.10 (0.48)	3.00 (0.78)	3.99 (0.84)	3.48 (0.66)
Religious	0.09 (2.89), 0.07	0.24 (1.37), 0.03	0.15 (2.07), 0.04	0.02 (5.29), 0.22
Yes	4.04 (0.47)	2.94 (0.80)	3.91 (0.86)	3.45 (0.66)
No	4.07 (0.51)	2.91 (0.80)	3.96 (0.83)	3.51 (0.63)
Area of residence	0.87 (0.03), 0.02	0.53 (0.40), 0.02	0.13 (2.36), 0.05	0.22 (1.50), 0.03
Rural	4.06 (0.49)	2.93 (0.79)	3.96 (0.83)	3.50 (0.63)
Urban	4.06 (0.52)	2.90 (0.80)	3.90 (0.87)	3.46 (0.66)
Blood donor	0.28 (1.17), 0.03	0.52 (0.41), 0.01	0.84 (0.04), 0.02	< 0.01 (6.36), 0.38
Yes	4.03 (0.56)	2.95 (0.82)	3.94 (0.79)	3.58 (0.66)
No	4.06 (0.49)	2.92 (0.80)	3.95 (0.85)	3.48 (0.64)
Organ donor	< 0.01 (11.07), 3.91	0.04 (4.20), 0.02	0.26 (1.28), 0.03	0.06 (3.64), 0.04
Yes	4.03 (0.56)	2.95 (0.82)	3.94 (0.79)	3.58 (0.66)
No	4.06 (0.49)	2.92 (0.80)	3.95 (0.85)	3.48 (0.64)

Note: The values in the table are either probability values (p) with Levene's test F-statistic (Levene'sF) and Bayes Factor (BF) or means (M) with Standarddeviations (SD). Bold=significant with $p < 0.05$.

Predicting motives for potential withdrawal

All of the respondents were participating in a biobank at the time of the study. However, this does not imply that their long-term participation was guaranteed. We therefore asked the respondents under what circumstances they would consider withdrawing from further participation. Accordingly, we performed linear multiple regressions to predict the importance of participants' motives for withdrawal, considering demographic variables, motives for participation, and prosocial psychological characteristics, while controlling for motives relating to either individual or societal harm (see Table 5). All assumptions for both multiple linear regression analyses were met: the normality p-plot and the scatterplot of standardised residuals showed that the data met the assumptions of normality, linearity and homogeneity of variance; there were no indications of multicollinearity (VIF scores < 10.0 ; Tolerance > 0.1). Significant predictors of withdrawal based on anticipated societal harm were: a stronger motivation to benefit society via participation ($B = 0.05$), a weaker motivation to benefit individually via participation ($B = -0.10$), being older ($B = 0.12$), being female ($B =$

0.06), better self-reported health ($B = 0.04$), being an organ donor ($B = 0.04$), a higher SVO score ($B = 0.04$), stronger biospheric values ($B = 0.08$), lower levels of societal trust ($B = -0.05$), a high level of trust in the government's DM&H ($B = 0.08$), and a low level of trust in commercial enterprises' DM&H ($B = -0.08$). The total explained variance was 19 %.

We conducted a similar analysis for individual harm considered as a motive for withdrawal, while controlling for societal harm as the motive of withdrawal. All assumptions were controlled and met, as previously described. To some extent, the results of this analysis mirrored the previous results. We found that being female ($B = 0.07$), having a partner ($B = 0.05$), a higher education status ($B = 0.07$), being a blood donor ($B = 0.05$), not being an organ donor ($B = -0.08$), a lower SVO score ($B = -0.04$), and a higher level of social trust ($B = 0.07$) were significant predictors of individual harm motives for withdrawal. There were no associations with the degree to which perceptions of individual or societal benefits motivated respondents' participation in the biobank. The total explained variance was 17 %.

Willingness of respondents to allow their data to be linked

We attempted to predict participants' intentions of sharing diverse personal data through linkages to medical research records according to participants' demographic variables, motives for participation and withdrawal, and prosocial psychological characteristics. All assumptions for this linear multiple regression analysis were met as well: the normality probability plot and the scatterplot of standardised residuals showed that the data met the assumptions of normality, linearity and homoscedasticity; there were no indications of multicollinearity (VIF scores < 10.0 ; Tolerance > 0.1). Several significant predictors of the participants' willingness to link their data to large-scale datasets for use in medical research were identified. These predictors were low levels of concern regarding potential harm relating to participation (B (Societal) = -0.08 , B (Individual) = -0.10), a strong motivation to realize benefits via participation (B (Societal) = 0.20 , B (Individual) = 0.05), a lower age ($B = -0.13$), being male ($B = -0.07$), a higher educational level ($B = 0.09$), being an organ donor ($B = 0.07$), a high level of trust in society ($B = 0.09$), trust in the government's DM&H ($B = 0.08$), trust in commercial enterprises' DM&H ($B = 0.14$), and trust in commercial researchers' DM&H ($B = 0.06$). The total explained variance was 21%.

Table 5. Results of the linear regression of motives for withdrawal and data linkages

Variable	Step	Predicting societal harm as motive of withdrawal			Predicting individual harm as motive of withdrawal			Predicting Intention to link personal data		
		Beta	95% CI unst. B	p	Beta	95% CI unst. B	p	Beta	95% CI unst. B	p
Withdrawal motive societal harm	1	-	-	-	0.36	[0.24, 0.30]	< 0.01	-0.08	[-0.10, -0.03]	< 0.01
Withdrawal motive individual harm	1	0.35	[0.41, 0.51]	< 0.01	-	-	-	-0.1	[-0.14, -0.06]	< 0.01
Participation motive societal benefits	2	0.05	[0.01, 0.15]	0.03	-0.04	[-0.11, 0.00]	0.06	0.2	[0.21, 0.32]	< 0.01
Participation motive individual benefits	2	-0.1	[-0.14, -0.06]	< 0.01	-0.04	[-0.07, 0.00]	0.06	0.05	[0.01, 0.08]	0.01
Age	3	0.12	[0.00, 0.01]	< 0.01	-0.04	[-0.01, 0.00]	0.18	-0.13	[-0.01, 0.00]	< 0.01
Gender (Male = 1, female = 2)	3	0.06	[0.03, 0.16]	< 0.01	0.07	[0.04, 0.14]	< 0.01	-0.07	[-0.14, -0.04]	< 0.01
Marital status (Without partner = 0, Partner = 1)	3	-0.03	[-0.14, 0.02]	0.16	0.05	[0.02, 0.14]	0.01	0.02	[-0.02, 0.10]	0.27
Educational level (Low = 1, Moderate = 2, High = 3)	3	-0.03	[-0.03, 0.00]	0.21	0.07	[0.01, 0.04]	< 0.01	0.09	[0.02, 0.05]	< 0.01
Paid job (no = 0, yes = 1)	3	0.02	[-0.05, 0.11]	0.5	0	[-0.07, 0.05]	0.79	-0.02	[-0.09, 0.04]	0.42
Religious (no = 0, yes = 1)	3	0.01	[-0.04, 0.09]	0.49	-0.03	[-0.09, 0.00]	0.1	-0.02	[-0.08, 0.02]	0.25
Residence (rural = 1 urban = 2)	3	-0.02	[-0.11, 0.04]	0.36	-0.02	[-0.09, 0.02]	0.21	0.02	[-0.03, 0.09]	0.29
Self-reported health	3	0.04	[0.00, 0.09]	0.05	-0.01	[-0.04, 0.03]	0.62	0.04	[-0.01, 0.07]	0.06
Frequency of charitable donations	4	0	[-0.03, 0.03]	0.96	0.02	[-0.01, 0.03]	0.47	0.03	[-0.01, 0.04]	0.14
Blood donor (no = 0, yes = 1)	4	-0.01	[-0.13, 0.06]	0.48	0.05	[0.02, 0.17]	0.02	0.03	[-0.01, 0.14]	0.11
Organ donor (no = 0, yes = 1)	4	0.04	[0.01, 0.14]	0.03	-0.08	[-0.16, -0.06]	< 0.01	0.07	[0.04, 0.14]	< 0.01
Social Value Orientation score	5	0.04	[0.00, 0.00]	0.05	-0.04	[0.00, 0.00]	0.05	-0.03	[0.00, 0.00]	0.16
Healthspenic values	6	0.03	[-0.03, 0.08]	0.41	0.02	[-0.03, 0.06]	0.53	-0.05	[-0.08, 0.00]	0.08
Hedonic values	6	0.02	[-0.03, 0.07]	0.36	0.02	[-0.02, 0.05]	0.39	0	[-0.04, 0.04]	0.9
Biospheric values	6	0.08	[0.02, 0.12]	< 0.01	-0.03	[-0.05, 0.02]	0.29	0.01	[-0.03, 0.05]	0.6
Societal trust	7	-0.05	[-0.12, -0.01]	0.02	0.07	[0.02, 0.11]	< 0.01	0.09	[0.05, 0.14]	< 0.01
TDM&H/government	7	0.08	[0.06, 0.20]	< 0.01	-0.04	[-0.10, 0.00]	0.08	0.08	[0.05, 0.16]	< 0.01
TDM&H/commercial enterprises	7	-0.08	[-0.14, -0.04]	< 0.01	-0.02	[-0.06, 0.02]	0.31	0.14	[0.09, 0.16]	< 0.01
TDM&H/commercial researchers	7	-0.03	[-0.08, 0.02]	0.22	0	[-0.04, 0.02]	0.74	0.06	[0.01, 0.09]	0.01
F (model 7)		25.3			21.02			27.54		
Adjusted R ²		0.19			0.17			0.21		

Note: B = regression coefficient, CI = confidence interval, Bold = $p < 0.05$.

Discussion

We investigated respondents' motives for participating and potentially withdrawing their participation in a large-scale, centralized data repository for supporting scientific research in the Netherlands. In addition, we investigated associations between their motives for participation and (potential) withdrawal, respectively, and their demographic and psychological characteristics as well as their willingness to permit their personal data to be linked with a population-based biobank. Our results indicate that participants' motives for both participation and potential withdrawal were clearly premised on societal and individual consequences. Whereas motives for participation were related to individual and societal benefits, motives for withdrawal were based on percep-

tions of potential societal or individual harm associated with participation in the biobank.

In line with previous studies, we found that both participation and (potential) withdrawal were primarily driven by societal motives [19,24]. However, our findings about participation and withdrawal, which were based on assumptions of individual benefit and harm respectively, suggest that individual motives for participation, especially those relating to potential harm, cannot be neglected. For example, highly educated participants were more likely to consider withdrawal because of potential individual harm than those who were less well-educated. Conversely, less well-educated participants were more likely to consider individual benefits when deciding to participate than highly educated participants. These findings confirm the results of a review study about motives to enrol in population-based biobank studies, which indicated the importance of consequences on an individual level [14]. Recent studies also found that predictors of individuals' willingness to contribute genetic data related to their positive evaluations of genetic data for the benefit of society and themselves [37,48].

In our model, trust was evidently a relevant predictor of motives for withdrawal. While these findings confirm those of previous studies regarding the importance of trust in biobank participation decisions [6,21,24], they also highlight the complexity of trust, especially in relation to withdrawal decisions. Anticipated societal harm as a motive for withdrawal associated with a lack of trust in society and in commercial enterprises and researchers, whereas anticipated harm to individuals motivating withdrawal decisions associated with a high level of trust in society. We might explain this counterintuitive finding as an indication of trust reducing perceived risks or anticipated harm for society when this trust is directed towards society including commercial enterprises. Individual harm motives, such as having a negative experience or cost of time, became more important when societal trust was high. In contrast, motives of anticipated societal harm were higher when trust is merely directed to government' data management and handling.

Although there is strong evidence that trust is positively associated with stronger intentions to participate [6,20], our findings support a view of trust as a conditional and fluid concept [6]. This is in line with recent results of an Anglo-Saxon study which showed a robust variability among trust categories in society, in particular the health care context [21]. Our results fit a definition of trust in society as a resource to cope with uncertainties [49]. Involvement of commercial enterprises in research, for example, is previously found to be a robust trigger for concerns of potential commodification or privacy violation which in combination with lack of trust led to withdrawal. These concerns touch upon

current developments in biobanking [18]. Broad societal trust, including trust in commercial enterprises or researchers, might be a resource to cope with uncertainties in participation. Narrow societal trust – only directed on government behaviour towards data – is a limited resource. A limited resource is likely to draw attention away from pragmatic concerns on an individual level. Accordingly, higher levels of trust do not necessarily hinder decisions to withdraw from participation.

Furthermore, we found that participants' motives were associated with several other psychological characteristics. Motives relating to societal benefits or societal harm were associated with stronger prosocial indicators, such as donor status, healthspheric or biospheric values, and SVO, confirming the findings of previous studies that highlighted the importance of general attitudes regarding perceptions of the risks and especially the benefits of participation in a large-scale, centralized data repository for medical research [14,33]. On top of that, several demographic factors, such as lower educational levels, being male, or not being an organ donor were predictors of stronger motives to pursue individual benefits via participation. By contrast, both individual and societal harm were important motives for women in potential decisions to withdraw participation. The same we found for blood donors in relation to individual harm, and in relation to societal harm motives for older participants and participants with a higher education status.

These findings suggest that gender, donor status, age, and education play different roles in participation and withdrawal in a biobanking context. Being female, blood donor, older or well-educated may increase the likelihood of an individual's withdrawal during changes, but these characteristics may also be positively linked to stronger perceptions of the benefits of medical research, greater comfort with giving blood, and trust in the biobank [24]. Our results partly support studies showing that females are stronger oriented towards relational or altruistic goals, while men are oriented on their own or own group goals in participation decisions [23,50]. Yet they emphasize results of a recent study about gender, age, and educational level differences, which showed that these differences associated inconsistently to volunteering when taking into account context or other demographic and prosocial intrapersonal characteristics, such as age, marital status, values, and societal trust [51]. An Australian study about biobanking confirmed these complexities in biobanking, since individuals with low educational levels are reportedly more willing to participate in biobanks that share resources or collaborate with commercial researchers than those who do not. Men were found to have more trust in biobanks collaborating with international researchers, while women had more trust in biobanks collaborat-

ing with national researchers [30]. Experimental designs in future studies seem to be worthwhile in light of these findings, since these designs, such as randomized controlled studies are better able to distinguish robust associations, and conditional associations between participation/non-participation and characteristics. For example, these designs could unravel the unique effects of blood and organ donor status.

Our study extends the focus of previous studies on key mechanisms of participation in biobanking through an investigation of the factors that influence the probability of individuals' withdrawal and their willingness to accept innovations in data collection, for example data linking. The drivers of participants' willingness to accept data linkages for medical research include a strong motivation to realize individual and especially societal benefits via participation, weak concerns about possible individual or societal harm, and high levels of societal trust and trust in the whole context of research. This finding supports those of previous studies indicating the importance of perceived benefits, risks, and trust [21,33]. In addition, we found greater willingness to accept data linkages among young, male, and well-educated participants. Some studies have found as well that men, younger people, and higher educated people are more confident about participation than women, older people and lower educated people, for example by perceiving fewer risks relating to genetic technologies [20,33], perceiving more genetic exceptionalism [20], or having a broad societal trust [21].

Using data linkage for enriching data repositories could provide new opportunities for successfully obtaining more data of young, male, and well-educated participants in biobanks. Further, data repositories might recruit more young people and men when offering data linkages possibilities as these groups are otherwise less likely than others to participate in research. The prospect of data linkage might be more attractive for them, because of the lower effort participation brings. Another reason could be that the potential to find interesting new information pertains to these individuals. However, our results also showed that women, older people, and those with a lower education status are less inclined to contribute their personal information via data linkages, thus indicating and confirming the need for tailored recruitment or retainment strategies based on these characteristics [36]. Our findings may indicate that new information technology solutions can only partly be a solution for recruitment and retainment of participants, as more data isn't synonymous with better research data. Further research could focus on investigating the factors that motivate these groups to be less or more inclined towards data-linking practices, and to unravel the variety of conditions for acceptance of data-linking practices.

More generally, whereas some groups within the population may be easier to recruit, other groups may be easier to retain when contextual changes in biobanking occur. These findings raise new questions about the current strict ideals on research ethics. For example, the ideal not to coerce individuals into participation in research with potential individual benefits, or to traverse the line of 'safe' contributions towards build-up of knowledge as a common good. As certain characteristics associate differently with motives and levels of trust depending on the context, a conservative approach on research ethics might have negative consequences for equal chances to participate in data repositories. Simultaneously, participation in data repositories might have more implications in the prospect of learning health systems [52], though the challenges of data security remain or are likely to increase [1].

A more progressive approach on reciprocity in biobank participation could accept and take into account populations' diversity in trust and values about benefits of research, especially for groups based on gender or age. By focusing on how to proactively deal with these differences in recruiting and retention of participation, it may be possible to prevent collisions of principles of research ethics in a fundamental way [18]. For example, prevention of potential discrimination of certain groups of individuals in policies seems to deny these individuals being appropriately represented in research. Concepts of equity or justice will become more salient in innovative data repository and information technologies which aim to optimize participation and data density rates [3]. As these are being built, it is primarily a societal question how to adapt to their uncertainties and effects [1]. That is why adaptive and sustainable systems for safeguarding research participants and non-participants from harm should concurrently be discussed and developed, for example with harm mitigation bodies [53]. These systems require to take the national and cultural context into account, such as levels of trust among a particular society or important values for populations in society [54]. By doing so, trust in these systems might be retained during uncertain contexts characterized by change, for example increasing concerns about potential commodification, privacy violations, and data security. The importance of the national and cultural context robustly explained differences in decision making about (non)participation both global [55] and within Europe [48]. Further research should aim to provide more evidence on the factors that increase participants' concerns and potential withdrawal in relation to their perceptions of impact of (non)participation on future policies or treatment based on research. Especially, interaction between different characteristics and motives merit further investigation.

Although our study yielded insights into the motives behind participation

and withdrawal of participation in biobanks, it had some limitations. All of the respondents were participants in a particular Dutch population-based biobank. Therefore, our findings may to some extent be specific to the Dutch societal context, population-based biobanking and prospective cohort studies. For example, distinctions between the public and private sectors may vary across countries as well as engagement in biobanking or concerns about privacy [6,48]. In addition to that, our participants group could differ substantially in demographic and prosocial characteristics from those who have not or not yet participated [25]. These differences may apply as well to samples of different types of biobanks and cohort studies. Furthermore, causal effects are difficult to capture through exploratory cross-sectional online surveys. This issue could be resolved by using more experimental designs in future studies and qualitative investigations of these processes. Although our findings were statistically reliable, the models' performances were relatively weak in explaining variance. Our instruments are tailored to our research question using partly validated measures and subscales, which might limit the generalisation and validity of our results. Thus, while our findings are illuminating to some extent, the mechanisms underlying participation and withdrawal require further elucidation. Nevertheless, our findings may provide valuable insights for developing effective recruitment strategies, data collection methods, or data repositories.

Conclusion

We have shown that motives for participation in and potential withdrawal of participation from a biobank can be differentiated into those relating to societal benefits and harm on the one hand and those relating to individual benefits and harm on the other hand. In addition, these motives are differently associated with participants' demographic and intrapersonal characteristics. Our data suggest that the tendency to withdraw from participation could be countered by inducing perceptions of more individual benefits in case of perceptions of limited societal benefits and low narrow levels of trust in society. Yet the results emphasize the complexity and potential trade-offs in perceived harms for others. Our findings may contribute to improving recruitment and retainment strategies for large-scale medical data collection and eliciting participants' agreement with data linkage by incorporating relevant values and highlighting the important benefits of the research for individuals.

References

1. O'Doherty KC, Christofides E, Yen J, Bentzen HB, Burke W, Hallowell N, et al. If you build it, they will come: unintended future uses of organised health data collections. *BMC Med Ethics* [Internet]. BMC Medical Ethics; 2016;17:16. Available from: <http://bmcmethics.biomedcentral.com/articles/10.1186/s12910-016-0137-x>
2. Rivera SM, Brothers KB, Cadigan RJ, Harrell HL, Rothstein MA, Sharp RR, et al. Modernizing Research Regulations Is Not Enough: It's Time to Think Outside the Regulatory Box. *Am J Bioeth.* 2017;17:1–3.
3. Mittelstadt BD, Floridi L. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Sci Eng Ethics.* Springer; 2016;22:303–41.
4. Stricker BH. Epidemiology and 'big data.' *Eur J Epidemiol* [Internet]. Springer Netherlands; 2017;32:535–6. Available from: <http://link.springer.com/10.1007/s10654-017-0294-3>
5. Kalkman S, Mostert M, Gerlinger C, van Delden JJM, van Thiel GJMW. Responsible data sharing in international health research: a systematic review of principles and norms. *BMC Med Ethics* [Internet]. BioMed Central; 2019 [cited 2019 Apr 11];20:21. Available from: <https://bmcmethics.biomedcentral.com/articles/10.1186/s12910-019-0359-9>
6. Gaskell G, Gottweis H, Starkbaum J, Gerber MM, Broerse J, Gottweis U, et al. Publics and biobanks: Pan-European diversity and the challenge of responsible innovation. *Eur J Hum Genet.* 2013;21:14–20.
7. Scholtens S, Smidt N, Swertz MA, Bakker SJLL, Dotinga A, Vonk JM, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. *Int J Epidemiol.* 2015;44:1172–80.
8. Hewitt R, Watson P. Defining biobank. *Biopreserv Biobank.* 2013;11:309–15.
9. Snell K, Tarkkala H. Questioning the rhetoric of a 'willing population' in Finnish biobanking. *Life Sci Soc Policy* [Internet]. Life Sciences, Society and Policy; 2019;15:1–11. Available from: <https://lssjournal.biomedcentral.com/articles/10.1186/s40504-019-0094-5>
10. Luiten A, Hox J, Leeuw E De. Survey Nonresponse Trends and Fieldwork Effort in the 21st Century : Results of an International Study across Countries and Surveys. 2020;36:469–87.
11. De Leeuw E, Hox J, Luiten A. International Nonresponse Trends across Countries and Years: An analysis of 36 years of Labour Force Survey data. *Surv Insights Methods from F* [Internet]. 2018;1–11. Available from: <http://data.aussda.at/dataverse/smif>
12. Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical

- instruments. *Genome Med* [Internet]. *Genome Medicine*; 2020;12:44. Available from: <https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-020-00742-5>
13. Jo A, Coronel BD, Coakes CE, Mainous AG. Is There a Benefit to Patients Using Wearable Devices Such as Fitbit or Health Apps on Mobiles? A Systematic Review. *Am J Med* [Internet]. Elsevier Inc.; 2019;132:1394–1400.e1. Available from: <https://doi.org/10.1016/j.amjmed.2019.06.018>
 14. Nobile H, Vermeulen E, Thys K, Bergmann MM, Borry P. Why do participants enroll in population biobank studies? A systematic literature review. *Expert Rev Mol Diagn* [Internet]. 2013;13:35–47. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23256702>
 15. Gayet-Ageron A, Rudaz S, Perneger T. Biobank attributes associated with higher patient participation: A randomized study. *Eur J Hum Genet* [Internet]. 2016;25:31–6. Available from: <http://www.nature.com/doi/10.1038/ejhg.2016.132>
 16. Ahram M, Othman A, Shahrouri M, Mustafa E. Factors influencing public participation in biobanking. *Eur J Hum Genet* [Internet]. Nature Publishing Group; 2014;22:445–51. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3953902&tool=pmcentrez&render-type=abstract>
 17. Johnsson L, Hansson MG, Eriksson S, Helgesson G. Patients' refusal to consent to storage and use of samples in Swedish biobanks: Cross sectional study. *Bmj* [Internet]. 2008;337:224–6. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2656925&tool=pmcentrez&rendertype=abstract>
 18. Ursin L, Ytterhus B, Christensen E, Skolbekken J-A. «If you give them your little finger, they'll tear off your entire arm»: losing trust in biobank research. *Med Heal Care Philos* [Internet]. Springer Netherlands; 2020;23:565–76. Available from: <https://doi.org/10.1007/s11019-020-09969-w>
 19. Kettis-Lindblad Å, Ring L, Viberth E, Hansson MG. Genetic research and donation of tissue samples to biobanks. What do potential sample donors in the Swedish general public think? *Eur J Public Health*. 2006;16:433–40.
 20. Middleton A, Milne R, Thorogood A, Kleiderman E, Niemiec E, Prainsack B, et al. Attitudes of publics who are unwilling to donate DNA data for research. *Eur J Med Genet* [Internet]. Elsevier; 2019;62:316–23. Available from: <https://doi.org/10.1016/j.ejmg.2018.11.014>
 21. Milne R, Morley KI, Howard H, Niemiec E, Nicol D, Critchley C, et al. Trust in genomic data sharing among members of the general public in the UK, USA, Canada and Australia. *Hum Genet* [Internet]. Springer Berlin

Heidelberg; 2019;138:1237–46. Available from: <https://doi.org/10.1007/s00439-019-02062-0>

22. Bossert S, Kahrass H, Strech D. The public's awareness of and attitude toward research biobanks - A regional German survey. *Front Genet.* 2018;9:1–11.
23. Richter G, Borzikowsky C, Lieb W, Schreiber S, Krawczak M, Buyx A. Patient views on research use of clinical data without consent: Legal, but also acceptable? *Eur J Hum Genet* [Internet]. Springer US; 2019;27:841–7. Available from: <http://dx.doi.org/10.1038/s41431-019-0340-6>
24. Critchley C, Nicol D, Otlowski M, Stranger M. Predicting intention to biobank: A national survey. *Eur J Public Health.* 2012;22:139–44.
25. Broekstra R, Aris-Meijer J, Maeckelberghe E, Stolk R, Otten S. Demographic and prosocial intrapersonal characteristics of biobank participants and refusers: the findings of a survey in the Netherlands. *Eur J Hum Genet* [Internet]. Springer US; 2021;29:11–9. Available from: <http://www.nature.com/articles/s41431-020-0701-1>
26. Critchley C, Nicol D, McWhirter R. Identifying public expectations of genetic biobanks. *Public Underst Sci* [Internet]. 2017;26:671–87. Available from: <http://pus.sagepub.com/cgi/doi/10.1177/0963662515623925>
27. Kaufman DJ, Murphy-Bollinger J, Scott J, Hudson KL. Public Opinion about the Importance of Privacy in Biobank Research. *Am J Hum Genet* [Internet]. The American Society of Human Genetics; 2009;85:643–54. Available from: <http://dx.doi.org/10.1016/j.ajhg.2009.10.002>
28. Skovgaard LL, Wadmann S, Hoeyer K. A review of attitudes towards the reuse of health data among people in the European Union: The primacy of purpose and the common good. *Health Policy (New York)* [Internet]. Elsevier Ireland Ltd; 2019;123:564–71. Available from: <https://doi.org/10.1016/j.healthpol.2019.03.012>
29. Aitken M, De St Jorre J, Pagliari C, Jepson R, Cunningham-Burley S. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. *BMC Med Ethics* [Internet]. BMC Medical Ethics; 2016;17:1–24. Available from: <http://dx.doi.org/10.1186/s12910-016-0153-x>
30. Critchley C, Nicol D, Otlowski M. The impact of commercialisation and genetic data sharing arrangements on public trust and the intention to participate in biobank research. *Public Health Genomics.* 2015;18:160–72.
31. Steinsbekk KS, Ursin LØ, Skolbekken JA, Solberg B. We're not in it for the money-lay people's moral intuitions on commercial use of "their" bio-

- bank. *Med Heal Care Philos* [Internet]. 2013;16:151–62. Available from: <http://dx.doi.org/10.1007/s11019-011-9353-9>
32. Nicol D, Critchley C, McWhirter R, Whitton T. Understanding public reactions to commercialization of biobanks and use of biobank resources. *Soc Sci Med* [Internet]. Elsevier Ltd; 2016;162:79–87. Available from: <http://dx.doi.org/10.1016/j.socscimed.2016.06.028>
33. Connor M, Siegrist M. Factors influencing people's acceptance of gene technology: The role of knowledge, health expectations, naturalness, and social trust. *Sci Commun*. 2010;32:514–38.
34. Siegrist M, Cvetkovich G. Perception of Hazards: The Role of Social Trust and Knowledge. *Risk Anal* [Internet]. Siegrist, Michael, Department of Psychology, University of Zurich, Plattenstrasse 14, 8032, Zurich, Switzerland: Blackwell Publishing; 2000;20:713–9. Available from: <http://search.ebscohost.com.proxy-ub.rug.nl/login.aspx?direct=true&db=psych&AN=2002-04285-005&site=ehost-live&scope=site>
35. Ridgeway JL, Han LC, Olson JE, Lackore KA, Koenig BA, Beebe TJ, et al. Potential Bias in the Bank: What Distinguishes Refusers, Nonresponders and Participants in a Clinic-Based Biobank? *Public Health Genomics* [Internet]. 2013;16:118–26. Available from: <http://www.karger.com?doi=10.1159/000349924>
36. Sanderson SC, Brothers KB, Mercaldo ND, Clayton EW, Antommaria AHM, Aufox SA, et al. Public Attitudes toward Consent and Data Sharing in Biobank Research: A Large Multi-site Experimental Survey in the US. *Am J Hum Genet*. 2017;100:414–27.
37. Middleton A, Milne R, Howard H, Niemiec E, Robarts L, Critchley C, et al. Members of the public in the USA, UK, Canada and Australia expressing genetic exceptionalism say they are more willing to donate genomic data. *Eur J Hum Genet* [Internet]. Springer US; 2019;424–34. Available from: <http://dx.doi.org/10.1038/s41431-019-0550-y>
38. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the LifeLines cohort study. *PLoS One*. 2015;10:1–12.
39. Merdad L, Aldakhil L, Gadi R, Assidi M, Saddick SY, Abuzenadah A, et al. Assessment of knowledge about biobanking among healthcare students and their willingness to donate biospecimens. *BMC Med Ethics* [Internet]. BMC Medical Ethics; 2017;18:32. Available from: <http://bmcmethics.biomedcentral.com/articles/10.1186/s12910-017-0195-8>
40. Messick DM, McClintock CG. Motivational bases of choice in experimental games. *J Exp Soc Psychol* [Internet]. 1968;4:1–25. Available from: [https://doi.org/10.1016/0022-1031\(68\)90046-2](https://doi.org/10.1016/0022-1031(68)90046-2)

41. Murphy RO, Ackermann KA, Handgraaf MJJ. Measuring Social Value Orientation. *Judgm Decis Mak*. 2011;6:771–81.
42. Schwartz SH, Cieciuch J, Vecchione M, Davidov E, Fischer R, Beierlein C, et al. Refining the theory of basic individual values. *J Pers Soc Psychol* [Internet]. 2012;103:663–88. Available from: <http://doi.apa.org/getdoi.cfm?doi=10.1037/a0029393>
43. van der Werff E, Steg L. The psychology of participation and interest in smart energy systems: Comparing the value-belief-norm theory and the value-identity-personal norm model. *Energy Res Soc Sci* [Internet]. Elsevier Ltd; 2016;22:107–14. Available from: <http://dx.doi.org/10.1016/j.erss.2016.08.022>
44. Bouman T, Steg L, Kiers HAL. Measuring Values in Environmental Research : A Test of an Environmental Portrait Value Questionnaire. *Front Psychol*. 2018;9:1–15.
45. Broekstra R, Aris-Meijer J, Maeckelberge E, Otten S, Stolk R, Maeckelberghe E, et al. Trust in Centralized Large-Scale Data Repository: A Qualitative Analysis. *J Empir Res Hum Res Ethics* [Internet]. 2020;15:365–78. Available from: <http://journals.sagepub.com/doi/10.1177/1556264619888365>
46. Mayer RC, Davis JH, Schoorman FD. An Integrative Model of Organizational Trust. *Acad Manag Rev* [Internet]. 1995;20:709–34. Available from: <http://www.jstor.org/stable/258792>
47. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp. Armonk, NY: IBM Corp.; 2017.
48. Voigt TH, Holtz V, Niemiec E, Howard HC, Middleton A, Prainsack B. Willingness to donate genomic and other medical data: results from Germany. *Eur J Hum Genet* [Internet]. Springer US; 2020;28:1000–9. Available from: <http://dx.doi.org/10.1038/s41431-020-0611-2>
49. Luhmann N. Trust and Power [Internet]. John Wiley & Sons Inc.; 1979. Available from: <https://books.google.nl/books?id=WZu4AAAAIAAJ>
50. Decker O, Winter M, Braehler E, Beutel M. Between commodification and altruism: Gender imbalance and attitudes towards organ donation. A representative survey of the German community. *J Gend Stud*. 2008;17:251–5.
51. Gil-Lacruz AI, Marcuello C, Saz-Gil MI. Gender differences in European volunteer rates. *J Gend Stud* [Internet]. Routledge; 2019;28:127–44. Available from: <https://doi.org/10.1080/09589236.2018.1441016>
52. Wouters RHP, van der Graaf R, Voest EE, Bredenoord AL. Learning health care systems: Highly needed but challenging. *Learn Heal Syst*. 2020;4:1–6.

53. McMahon A, Buyx A, Prainsack B. Big Data Governance Needs More Collective Responsibility: The Role of Harm Mitigation in the Governance of Data Use in Medicine and Beyond. *Med Law Rev.* 2020;28:155–82.
54. Gaskell G, Stares S, Allansdottir A, Allum N, Castro P, Esmer Y, et al. Europeans and Biotechnology in 2010 - Winds of change? *Luxemb. Eur. Comm.* 2010.
55. Sheikh Z, Hoeyer K. “That is why I have trust”: unpacking what ‘trust’ means to participants in international genetic research in Pakistan and Denmark. *Med Heal Care Philos.* Springer Netherlands; 2018;21:169–79.

Chapter 6

Dutch cohort studies and social capital: an examination of trust and participation indicators

Abstract

Participation in public goods remains an ongoing challenge for societies. Cohort studies are no exception to this quest for receiving sufficient contributions, as their value for scientific research, policy research, or clinical care depends on full participation. As participation in research can be perceived as collective or prosocial behavior, being participant in a cohort study might imply a sense of connectedness towards members of society or lack of it. This paper therefore explores the relationship between social capital and participation in two specific public goods, being two Dutch cohort studies. We used unique cohort data about social capital from Statistics Netherlands linked with the Lifelines biobank and longitudinal cohort study. We assumed that people with more social capital are also more likely to be a participant in Lifelines. The relevance of social capital for participation in Lifelines corresponds with the expectations expressed in our hypothesis. We found that if a person is endowed with more social capital, i.e. greater levels of participation and trust, the likelihood of them being a participant in Lifelines is higher. Our findings confirm that social capital increases the probability that a person participates in multiple longitudinal cohort studies, as examples of participation across public goods. Social capital seems to be more than a prerequisite for participation, especially indicators of trust. Future research should investigate how these factors might be addressed to reduce biases in cohort studies.

Keywords: Social capital; Trust; Participation; cohort studies; biobank.

This chapter is submitted to Social Indicators Research on behalf of Broekstra R., Exel J., Aris-Meijer J.L., Maeckelberghe, E.L.M. Sabine Otten, R.P. Stolck, A.J.M.M. Luiten, HJ.J.G. Schmeets and currently under review.

Introduction

Initial and continuous participation in systems or repositories with a public purpose, known as ‘public goods’, remains an ongoing challenge for societies. Public goods rely in whole or in part on the contribution of citizens to be of value to society; however, such participation may have a short-term negative impact on the individual and therefore be a disincentive to contribute (Parks et al., 2013). This challenge to participation resonates in the quest for sufficient contributions of personal data, biosamples, blood or organs to enable scientific research, policy research or clinical care. For example, focusing on the context of research, we see a steady decrease in response rates in survey studies over the last 35 years (De Leeuw et al., 2018; Luiten et al., 2020). Moreover, the demand for organ donors in Europe is overall still higher than the number of registered organ donors, which indicates similar issues in the clinical care context (Caplan, 2016). The lack of participation in public goods indicates that society could benefit from a better understanding of a multitude of dimensions relevant to people’s decision-making about participation in public goods, especially in the domain of biomedical research.

There is ample empirical evidence that connectedness to society is related to compliance to all kinds of participation requests and that different participation behaviours are associated with each other. For example, in 1993, Brehm demonstrated that voting in the United States national elections positively correlated with survey participation. Furthermore, in the Netherlands, it was recently shown that organ donation is positively related to participation in the survey on Social Cohesion and Wellbeing (S&W survey) (Schmeets & Peters, 2021). In the same vein, studies have found a variety of similar indicators of connectedness to society, such as higher levels of societal trust, organ donation and prosocial orientation, which are related to participation in a large-scale, centralized data repository for scientific research (Broekstra et al., 2021; Critchley et al., 2012; Gaskell et al., 2013; Voigt et al., 2020).

Despite these indications that a better understanding of why people feel a connectedness to society could lead to a broader understanding of why they participate in public goods, little is known about how this connectedness relates to such participation. In the context of participation in scientific research, this understanding could be of particular importance, as there are increasing numbers of similar yet competing public goods, such as various large-scale cohort studies and data repositories. Moreover, there are some indications of individual, group and regional differences in participation across public goods. For ex-

ample, Northern European countries have higher participation rates in biomedical research than Southern European countries (Gaskell et al., 2013); however, Southern European countries have the largest proportion of registered organ donors in Europe (Matesanz et al., 2017).

Similar to organ donation, participation in scientific research can be perceived as a form of collective action (e.g. Healy, 2004; 2005; Ladin et al., 2015). Such participation is defined as cooperation among citizens to achieve a common goal, such as gaining knowledge through research. However, achieving this common goal of research has short-term individual costs for individual participants, such as investing time and effort, or the potential risk to their privacy, which might be a barrier to participation. A slightly alternative approach would be to define participation as a form of prosocial behaviour, as benefitting other individuals or groups (e.g. Bekkers, 2006; Merz et al., 2017; Broekstra et al., 2021; Piersma et al., 2019). In this definition, the achievement of a common goal is not an important condition, since individuals participate based on the motive to support others.

The dilemma of individual cost versus collective benefit for participants remains unresolved in this latter definition, although it takes a slightly altered form. The prosocial behaviour of participation should not lead to harm to oneself. Moreover, the total sum of benefits of the research to others depends on all other decisions about participation by those citizens invited to take part. While a certain goal does not have to be reached, the effect of prosocial behaviour through participation in a public good might be limited when few people contribute.

Despite the differences between the approaches, connectedness is a key element which the concept of social capital embodies in both. Putnam (2000, p. 19) defined social capital as ‘connections among individual social networks and the norms of reciprocity and trustworthiness that arise from them’. This definition comes very close to that adopted by the OECD and many institutes of statistics, including Statistics Netherlands, which define social capital as: ‘networks together with shared norms, values and understandings that facilitate co-operation within or among groups’ (Cote and Healy 2001, p. 41; see also Keeley 2007, pp. 102–105).

Relating this concept to connectedness to society as a prerequisite for participation in public goods such as scientific research, it can be argued that social capital may itself be a prerequisite for participation, especially participation in various scientific data repositories or cohort studies. Moreover, higher levels of social capital at the individual level may further encourage the sharing of information through participation in a specific public good due to the greater influ-

ence of social networks with shared values. Furthermore, it may positively relate to participation as a form of collective action (Kim, 2018). In contrast, people and communities endowed with distrust in others and in societal institutions, and who do not actively participate in society, have been found to be reluctant to participate in scientific research or research by governmental bodies (Broekstra et al., 2021).

Therefore, this paper will examine the relationship between social capital and participation in public goods, looking in particular at cohort studies that create scientific data repositories. Our aim is to advance our understanding of a multitude of dimensions relevant to people's decision to participate in public goods in the biomedical domain. We will undertake a detailed examination of the association between social capital indicators and participation in multiple cohorts studies, using a rich and unique dataset from the S&W survey for the period 2012–2018, containing data on more than 53,000 individuals aged 15 and over, and linked to data retrieved from the Lifelines biobank and cohort study in the northern Netherlands that has more than 167,000 participants.

Methods

Procedure and data

We compared the characteristics of respondents participating in a survey study (A) with respondents participating in an additional independent cohort study (B). The first study (A) required individual one-off participation in the S&W survey, which was conducted by Statistics Netherlands among a randomized sample of the Dutch population. The second study (B) requires longitudinal participation in the Lifelines biobank and cohort study about healthy ageing and is a pseudo-randomized sample.

The S&W survey was undertaken in the period 2012–2018 and includes over 53,000 individuals aged 15 and above. The data was collected in a sequential mixed-mode design. Based on a random sample selection, individuals were invited to participate in the survey by internet. After two reminders, the non-respondents were approached for a telephone or a face-to-face interview at home. Response rates varied from 63% to 70% in the 2012–2018 period. Individuals in the sample are weighted to reflect the composition of the population and reduce non-response bias. The survey contains social capital indicators, such as social contacts, volunteering, political participation, generalized trust and trust in institutions in the public and private domains. It also captures information on personal wellbeing, involvement in religion, education level and other background characteristics, partly retrieved from administrative data sources such as the

municipal population registries and the tax office (on factors such as income).

The Lifelines biobank and cohort study provides key infrastructure for biomedical research in the Netherlands. Like other biobanks, it stores large quantities of biological specimens as well as data extracted from questionnaires and measurements from healthy citizens, thereby aiming to facilitate studies on patients with specific diseases or prospective studies on the onset and progress of chronic diseases (Krokstad et al., 2013; Scholtens et al., 2015; Sudlow et al., 2015; UK Biobank Coordinating Centre, 2007; Van Staa et al., 2016). All 167,729 Lifelines participants consented to participate in the biobank between 2006 and 2013, allowing it to examine 10% of the population in the northern Netherlands. Lifelines applies a broad range of investigative procedures to this multigenerational cohort to assess biomedical, sociodemographic, behavioural, physical and psychological factors, with a particular focus on multi-morbidity and complex genetics. The Lifelines biosamples and data are donated at recruitment and follow-up sessions every one to five years, and are retained for a period of 30 years (Klijs et al., 2015; Scholtens et al., 2015). The participants are recruited through General Practitioners, participating family members and self registration. The response rate after invitation by General Practitioners was approximately 25% (Scholtens et al., 2015).

When comparing these cohort studies, participation in the Lifelines study differs from participation in the S&W survey in two ways: (1) Lifelines is not based on a random sample and (2) investment in a longitudinal cohort study is greater compared to a one-off 20 to 25 minute survey.

We linked the S&W survey data with the relevant Lifelines data to examine differences between respondents from the northern Netherlands who only participated in the S&W survey and respondents who participated in both cohort studies. This linkage of data was done after all necessary organizational approvals from Lifelines and Statistics Netherlands through a standard data linkage protocol and exchange servers at Statistics Netherlands. This allowed us to add an extra variable to the S&W survey data, which distinguished respondents participating in both studies, thus enabling us to examine the relationship between individual social capital and participation in Lifelines. By linking the 167,729 participants in the Lifelines cohort study with the 5,663 participants in the S&W survey living in the northern Netherlands, we were able to divide the S&W participants into two groups: (1) no additional participation in Lifelines ($n = 4,993$) and (2) additional participation in Lifelines ($n = 730$).

We also controlled our sample for multiple participation across surveys by Statistics Netherlands, although respondents in Statistic Netherlands samples are only allowed to participate once every 5 to 7 years in surveys by Statistics

Netherlands. Conforming to our expectations, we did not find respondents who had simultaneously participated in multiple surveys by Statistics Netherlands.

Social capital: concept and measurement

The concept of social capital was first introduced by Hanifan (1920), but it took more than 60 years before it became popular through the work of Bourdieu (1983) and Coleman (1988). Since then, many definitions, conceptualizations and measurements have been proposed in the literature (e.g. Grootaert & Narayan, 2004; Hudson, 2006; Isakka, 2006; Iyver, Kitson & Toh, 2005; Knack & Keefer, 1997; Lochner et al., 1999; Nieminen et al., 2008; Neira, Vazquez & Portela, 2009; Nurul Islam et al., 2011; Nurul Islam, Yew & Viswanatha, 2014; Paxton, 1999; Poortinga, 2006; Ponthieux, 2004; Portela, Neira & Del Mar Salinas-Jiménez, 2013; Portes, 1998; Putnam, 1995a; 2000; Stiglitz, Sen & Fitoussi, 2009). However, these approaches generally agree that social capital includes ‘social networks’ and ‘trust’, or they use trust and social networks as independent indicators (e.g. Hooghe & Vanhoutte, 2011; Paxton, 1999). For some, social capital is primarily understood as a community property (Neira, Vazquez & Portela, 2009). For others, it is also treated at the individual level, as possessed by individuals within the community (Gannon & Roberts, 2018). The latter approach to individual social capital is used in this study, since it fits both our assumption of the key role of intrapersonal characteristics in behaviour as well as the data available from the S&W survey.

Social capital in the Netherlands

The concept of social capital was first introduced by Hanifan (1920), but it took more than 60 years before it became popular through the work of Bourdieu (1983) and Coleman (1988). Since then, many definitions, conceptualizations and measurements have been proposed in the literature (e.g. Grootaert & Narayan, 2004; Hudson, 2006; Iisakka, 2006; Iyver, Kitson & Toh, 2005; Knack & Keefer, 1997; Lochner et al., 1999; Nieminen et al., 2008; Neira, Vazquez & Portela, 2009; Nurul Islam et al., 2011; Nurul Islam, Yew & Viswanatha, 2014; Paxton, 1999; Poortinga, 2006; Ponthieux, 2004; Portela, Neira & Del Mar Salinas-Jiménez, 2013; Portes, 1998; Putnam, 1995a; 2000; Stiglitz, Sen & Fitoussi, 2009). However, these approaches generally agree that social capital includes 'social networks' and 'trust', or they use trust and social networks as independent indicators (e.g. Hooghe & Vanhoutte, 2011; Paxton, 1999). For some, social capital is primarily understood as a community property (Neira, Vazquez & Portela, 2009). For others, it is also treated at the individual level, as possessed by individuals within the community (Gannon & Roberts, 2018). The latter approach to individual social capital is used in this study, since it fits both our assumption of the key role of intrapersonal characteristics in behaviour as well as the data available from the S&W survey.

Social capital composite index measure

In the literature, social capital is often based on a composite score in which a number of national trust and participation averages are included (e.g. Putnam, 2000, p. 291). Those studies measuring social capital on the individual level based on micro data often only employ a limited number of indicators of social capital, such as volunteering (Alesina & La Ferrara, 2000) and trust in other people, referred to as generalized trust (Rocco & Fumageli, 2014). In this study, we adopt the measure of individual social capital introduced by Statistics Netherlands (Van Beuningen & Schmeets, 2013), which is based on the social cohesion framework developed there in 2008. This framework measures social capital along the two key dimensions of participation and trust but – partly based on Putnam's social capital index (Putnam, 2000) – integrates them into a single index (Van Beuningen & Schmeets, 2013). This composite index includes nine indicators of a person's participation in society and eight indicators of trust (Appendix 1). For both the participation and trust dimensions, the model makes distinctions at the social (micro), organizational (meso) and political (macro) levels (see also Halpern, 2005; Sharp & Randhawa, 2012).

In terms of the dimension of participation, social participation captures contacts with family, friends, neighbours and help provided to other people (social level). It refers to social contacts with relevant others and is based on the hypothesis that potential resources are made available through social networks. Social contacts facilitate information flows and also the extent to which

people can receive or give support and information (Flap & Völker, 2004).

While social participation reflects informal activities, organizational participation is connected to formal activities within organizations in which people are connected to each other. It includes volunteering, activities in associations for leisure, sport, culture and the neighbourhood, and paid work (organizational level). The final level relates to politics. Following Putnam (1995a), political involvement, such as voting and taking political action, is found to be a prerequisite for democracy. It further creates connections with others supporting specific political parties and/or governmental policies. It is measured by voting in parliamentary elections and participation in other political activities (political level).

The second dimension of social capital is trust, which is necessary to create generalized reciprocity within a society. Social relations and formal activities are only useful if they are trusting and positive (Paxton, 1999). The trust dimension also has three levels, concerning information on trust at the social level; trust at the organizational level; and trust at the political level. Social trust, or generalized trust, refers to trust in other people and is a condition for forming positive, reciprocal ties with other people, and as such it increases willingness to act in favour of the community, which is important for building social capital (Putnam et al., 1993). At the organizational level, we can distinguish trust in institutions in the public domain, such as the police, lawyers, the army, civil servants and the media, as well as trust in large companies in the private domain. Finally, trust at the political level concerns the political institutions which ensure that civilians can function effectively in society, which in this study, includes trust in the Dutch parliamentary process.¹

Based on these three levels for each of the two dimensions, the social capital model thus contains six subdimensions, for which we used a total of 17 different indicators (see Appendix 2 for descriptive statistics). These consist of both behavioural as well as attitudinal indicators. Social participation was measured as the frequency of contacts with family, friends and neighbours, as well as helping other people in an informal manner at least once in four weeks. Participation in organizations was measured as active participation in associations at least once a month, having a paid job for at least 12 hours a week, and volunteering at least once a year. Political participation was operationalized as voting during the previous parliamentary elections and undertaking political activities in the last 5 years. Social trust was measured by asking whether the respondent trusts other people in general. Institutional trust was measured

¹ The various social as well as organizational participation items have been included in the Permanent Survey on Living Conditions (POLS) since 1997. The political participation items as well as the trust items are based on the Dutch Parliamentary Election Study (DPES).

with single questions on trust in the public domain (army, police, judges, civil servants, press) and the private domain (large companies) on a scale from 'not at all' to 'a lot'. Finally, political trust was measured as trust in the Dutch parliamentary process. Most questions included the option to 'decline to answer' and/or it was not mandatory to provide an answer ('no answer' option). This concerned all trust variables and, apart from volunteering, all questions concerning social participation.

We followed previous decisions in the treatment of missing observations (Van Beuningen & Schmeets, 2013), with missing items imputed by the median values. Furthermore, for some respondents, such as for the 15-17 age group and for people residing in the Netherlands who are not citizens, voting in parliamentary elections was not possible. For these groups, we assumed that not being eligible to vote would mean that they felt less connected to the Dutch national elections than otherwise, and therefore we imputed non-eligibility as 'not voted'. For engaging in paid work, we stipulated at least 12 hours per week to ensure that there were contacts at the workplace for more than one day per week.

We also based our construction of the social capital index on that introduced by Van Beuningen & Schmeets (2013), who showed that a formative model should be applied to this index rather than a traditional reflective model such as traditional factor analysis (Grootaert & Narayan, 2004). The formative model entailed Multiple Table Analysis, also known as a hierarchical components model (Lohmöller, 1989; Tenenhaus, 2004), and estimated in R using the PLSPM package (Wetzels, Odekerken-Schröder & Van Oppen, 2009). A path weighting scheme was used such that weights could be interpreted as regression coefficients. The model calculates the joint contribution of the 17 indicators to the overall social capital score.²

Kowalski (1972) notes that correlations can be biased due to non-normal distributions, which was the case for most indicators in this study. Appendix 3 presents the correlation matrix with the 17 indicators for the 2012-2018 period of the S&W survey. It is notable that correlations between the indicators range from -0.05 to 0.53, which is rather low. Moreover, the correlations are rather stable for the specific years. Thus, the correlation coefficient may not be an informative statistic of association. However, this non-normality does not violate the model's assumptions, since the PLS method is able to cope with highly skewed variables.

² Initially, the formative model was applied to the data of the Permanent Survey on Living Conditions (POLS) in 2010 (Van Beuningen and Schmeets, 2013). After a redesign of POLS, it was renamed the Survey on Social Cohesion and Wellbeing in 2012. This model was also applied for the consecutive years of 2012 to 2018.

Control measures: background characteristics, religious involvement and health status

We controlled our measurement of social capital for other mechanisms that are linked to participation in a public good. First, there are discrepancies in scores on individual social capital between subpopulations in the Netherlands. In particular, the less educated, the elderly (aged 65+) and non-Western or other immigrants score lower than the higher educated, younger age groups and native Dutch (Van Beuningen & Schmeets, 2015). In addition, social capital increases as income becomes higher. Marital status also matters, with social capital higher among married and single people than divorced and widowed people. Under the assumption that social capital is relevant to participation in public goods, these patterns should be reflected in the proportion of individuals who participate in the Lifelines cohort studies.

Along with the social-demographic background characteristics, previous studies have found other characteristics that should be included as controls in our models. Given that religion, and in particular attending religious services, is often considered to foster social capital it is important that we include it (for the US and Canada, see Putnam & Campbell, 2012; Smidt, 2003; Wuthnow, 2002; for the Netherlands, see Bekkers & Schuyt, 2008; De Hart, 2014; Schmeets, 2019). The survey contains data on whether an individual considers themselves to be part of a religious denomination. This question had the following answer categories: (1) No religion; (2) Roman Catholic; (3) Reformed; (4) Calvinist; (5) Dutch Protestant; (6) Islam; (7) Hindu; (8) Buddhist; (9) Jewish; (10) Other. Due to the limited number of Muslims, Hindus, Buddhists and Jewish people in the northern Netherlands, we aggregate these groups into the category 'Other'. The extent to which individuals practise their beliefs was determined with a question about how often an individual attends religious services, with the following answer categories: (1) At least once per week; (2) 2-3 times per month; (3) Once per month; (4) Less than once per month; (5) Rarely or never.

In addition to data on religion, we used data on level of education from the survey, as well as data on gender, age, marital status and migration background from municipal population registers. Another control variable was political interest, as the level of social capital is higher among people who are more involved in politics, and we expected that they would also be more willing to participate in Lifelines. Finally, we included subjective health status in the models, as a recent study indicated that self-reported healthier people are more motivated to participate in Lifelines (Broekstra et al., 2021).

Results

Table 1 presents the distribution of the auxiliary variables, the control variables and the substantive variable of social capital for all participants in the S&W survey from the northern Dutch population ($n = 5,663$), detailed by the share that is not linked ($n = 4,933$) and that is linked to Lifelines ($n = 730$). Overall, 13% of the people who participated in the S&W survey in the northern Netherlands also participated in Lifelines.

The social capital score was directly retrieved from the model. Before running the model, the missing observations (missing items) were computed by the median values. A linear transformation was used for the index scores, with a mean value of 0.0 and a standard deviation of 1.0 based on the yearly national sample in the 2012-2018 period.³ The mean value of the social capital composite index score was 0.0 and the standard deviation 1.0. The mean value (-0.0001) and standard deviation (0.987) hardly changed for the northern part of the country. The scores ranged from -4.3 to +2.7. Based on these scores, we constructed quintiles⁴, with the results for five social capital groups showing that an individual's social capital was related to participation in both studies: the higher the level of social capital, the greater the willingness to participate in both the S&W survey and in Lifelines. On average, 13% participated in the S&W survey and the Lifelines cohort studies (Table 1). People scoring lowest on the social capital index had an 8% participation rate in Lifelines. This percentage gradually increased to 19% for the highest social capital quintile.

3 For more technical details and various robustness checks, see Van Beuningen and Schmeets (2013).

4 The choice of five groups (quintiles) was mainly based on it allowing the possibility to include interactions with the controls in the models. As we needed sufficient statistical power, we created five rather than ten groups (e.g. based on percentiles). As a robustness check, we also used percentiles ranging from -1.36 (first) to 1.20 (last). This did not lead to other findings in comparison to the quintiles.

Table 1. - Descriptive statistics social cohesion and wellbeing
2012/2018 in the northern Netherlands

	Not participating in Lifelines		Participating in Lifelines		Total
	N	%	N	%	N
Total	4,933	87.1	730	12.9	5,663
Social capital index score					
1 (lowest quintile)	981	91.7	89	8.3	1,070
2	950	88.8	120	11.2	1,070
3	939	87.7	132	12.3	1,071
4	924	86.4	146	13.6	1,070
5 (highest quintile)	866	80.9	204	19.1	1,070
Gender					
Male	2,456	89	305	11	2,761
Female	2,477	85.4	425	14.6	2,902
Age					
15-24	753	95.2	38	4.8	791
25-34	596	85.3	103	14.7	699
35-44	680	83.7	132	16.3	812
45-54	771	76.6	236	23.4	1,007
55-64	906	88.3	120	11.7	1,026
65-74	787	90.5	83	9.5	870
75+	440	96.1	18	3.9	458
Marital status					
Married	2,573	84.9	457	15.1	3,030
Divorced	369	84.2	69	15.8	438
Widowed	293	93.6	20	6.4	313
Single	1,698	90.2	184	9.8	1,882
Education					
Elementary	435	92.2	37	7.8	472
Low	1,202	91.2	116	8.8	1,318
Middle	2,029	85.8	335	14.2	2,364
High	855	83.2	173	16.8	1,028
Highest	332	83.8	64	16.2	396

Standardised household income

First quartile (lowest)	953	92.6	76	7.4	1,029
Second quartile	1,304	89.7	150	10.3	1,454
Third quartile	1,407	84.4	261	15.6	1,668
Fourth quartile (highest)	1,246	83.7	243	16.3	1,489

Migration background

Dutch native	4,515	86.7	693	13.3	5,208
Migrant	418	91.9	37	8.1	455

Subjective health

Very bad / bad	245	92.1	21	7.9	266
Not bad / not good	1,072	90.2	116	9.8	1,188
Good / very good	3,564	85.8	589	14.2	4,153

Membership of associations

No	896	92.3	75	7.7	971
Yes	3,197	85.5	540	14.5	3,737

Interested in politics

Not at all	963	89.5	113	10.5	1,076
Little	1,407	85.9	230	14.1	1,637
Quite	1,963	86.4	309	13.6	2,272
A lot	464	87.5	66	12.5	530

Denomination

No religion	2,547	87.5	365	12.5	2,912
Roman Catholic	268	87.6	38	12.4	306
Reformed	372	85.9	61	14.1	433
Calvinist	326	85.1	57	14.9	383
Dutch Protestant	421	84.0	80	16.0	501
Other religion	291	90.9	29	9.1	320

Attendance religious services

Rarely/never	3,312	87.5	471	12.5	3,783
Less than once a month	175	82.2	38	17.8	213
Once a month	105	82.7	22	17.3	127
Two/three times a month	190	83.3	38	16.7	228
At least once a week	400	87.3	58	12.7	458

We also observed differences among the control variables of gender, age and marital status. First, we found that women participated more often in Lifelines than men. Moreover, individuals between 25 and 55 participated more in Lifelines than both the young and elderly, who showed lower levels of participation in Lifelines. Finally, we found lower numbers of Lifelines participants among widowed and single people compared to married or divorced people.

In addition, based on the descriptive statistics, education, income, migration background, health status and interest in politics were significantly, although weakly, associated with being a Lifelines participant. First, a higher level of education and income implied that people were more likely to be a Lifelines participant. Second, more Lifelines participants were found among those with a Dutch background than migrants. Third, people participating in the Lifelines cohort were more healthy and more interested in politics.

The results also revealed discrepancies in the willingness to participate in Lifelines between those who practise a religion and those who do not to varying degrees. Individuals who did not belong to a religious group or rarely attended religious services were less often participants in Lifelines compared to those who regularly attended religious services. However, the group that attended services most often – at least once per week – did not differ from the non-religious group in their degree of Lifelines participation. Furthermore, religious affiliation was not related to the propensity to become a Lifelines participant.

Based on these descriptive statistics, no firm conclusions can be drawn regarding the impact of social capital, religious involvement and the other variables on Lifelines participation, given the interdependence of relevant characteristics. To analyse the data in greater detail, we performed a multiple logistic regression analysis (Table 2). We gradually added variables to the model in order to illustrate how characteristics were interconnected, and to reveal their specific relevance to participation in one study (S&W) or in two studies (S&W and Lifelines).

Table 2. Logistic regression of participation in Lifelines in the northern Netherlands

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6			
	Wald	Sig.	O.R.	Wald	Sig.	O.R.	Wald	Sig.	O.R.	Wald	Sig.	O.R.		
Social capital index score (lowest quintile) (ref.)	57.9	***		18	**	11.3	*	15.2	**	17	**	10.2	*	
	2	*	1.39		1.19		1.19		1.15		1.17		1.18	
	3	**	1.55		1.26		1.24		1.19		1.24		1.18	
	4	***	1.74	†	1.34	†	1.32		1.28	†	1.33		1.29	
5 (highest quintile)	***	2.6	***	1.78	**	1.67	***		1.69	***	1.77	**	1.63	
Gender														
Male (ref.)														
Female	17.1	***	1.42	15.4	***	1.44	16.9	***	1.43	14.7	***	1.4	13.1	
Age	112.7	***		101.4	***		110.9	***		108.1	***		96.9	
15-24 (ref.)														
25-34	***		2.69	***		3.08	***	***	2.73	***		2.64	***	3.09
35-44	***		2.82	***		3.1	***	***	2.9	***		2.84	***	3.2
45-54	***		4.12	***		4.79	***	***	4.28	***		4.06	***	4.84
55-64	*		1.75	*		1.93	**	**	1.82	*		1.75	*	1.96
65-74	†		1.57	*		1.75	*	*	1.65	†		1.55	*	1.8
75+			0.71			0.89			0.77			0.71		0.94
Marital status	6.7	†		8.1	*		6.3	†		6.4	†		7.3	†
Married (ref.)														
Divorced			1.11			1.26			1.13			1.12		1.26
Widowed			0.9			0.83			0.92			0.91		0.86
Single	*		0.76	*		0.76	*	*	0.77	*		0.76	*	0.77
Education														
Elementary (ref.)	7.1			5.9			7.2			7.1			6.6	
Low			0.8			0.73			0.77			0.76		0.69
Middle			1.09			1.01			1.05			1.04		0.97
High			1.05			0.93			1.02			1.01		0.91
Highest			0.88			0.88			0.85			0.86		0.87

Standardised income household									
First (lowest) quartile (ref.)	23	***	20.7	***	21.8	***	23.5	***	20.5
Second quartile	†	1.33	†	1.33	†	1.33	†	1.36	†
Third quartile	***	1.86	***	1.89	***	1.84	***	1.9	***
Fourth (highest) quartile	***	1.84	***	1.86	***	1.81	***	1.87	***
Migration background									
Dutch native (ref.)	4.9	*	0.66	5.9	*	0.6	5.3	*	0.64
Migrant							4.1	*	0.68
Attendance religious services									
Rarely/never			7.5						7.2
Less than once a month			†	1.4					1.39
Once a month			†	1.53					†
Two/three times a month				1.36					1.36
At least once a week				1.2					1.16
Subjective health / health status									
Very bad / bad				4.9	†				1.3
Not bad / not good						1.01			0.88
Good / very good						1.29			1.02
Interested in politics									
Not at all							3		2.5
Little								1.17	1.13
Quite								1	0.95
A lot								0.98	0.95
Nagelkerke R square (%)									
Missing	2	10	10.7		10.2		9.9		10.7
N	312	410	1,192		443		517		1,272
	5,351	5,253	4,471		5,220		5,146		4,391

Note: † p < 0.10 * p < 0.05 ** p < 0.01 *** p < 0.001

Model 1 in Table 2 shows a pattern that is strongly reminiscent of the descriptive statistics. We observe a gradual increase in Lifelines participation as individual social capital increases. This indicates that people who are more involved in society, that is, have higher participation and trust rates, are also more likely to participate in the Lifelines study. We observe a gradual increase in the odds of participating in Lifelines as social capital increases (odds ratios: 1.4; 1.5; 1.7 and 2.6). The odds of a person who scored highest on social capital (highest quintile) being a Lifelines participant is almost three times higher than a person with low social capital (lowest quintile). Note that the impact remains linear: each addition to social capital equally increases the chance of being a Lifelines participant.

Model 2 includes gender, age, marital status, education, income and migration background. When controlling for these characteristics, the impact of social capital only slightly decreases; thus, there remains strong evidence that an individual's social capital increases the propensity to become a Lifelines participant in addition to being a participant in the S&W survey. Some of these characteristics had an independent effect on the probability of being a Lifelines participant. Women were 43% more likely to be a Lifelines participant than men. Age shows that the probability of being a Lifelines participant increased after 25, remained relatively stable until age 45, peaked for the 45-54 age group, and subsequently levelled out. The higher the income, the more likely one is to be a member of Lifelines. More specifically, we can observe that people living in a household with above average income (third and fourth quartiles) are more often participants in Lifelines than those in lower income groups. Furthermore, migrants are less often Lifelines respondents than native Dutch. Finally, and not in accordance with the descriptive statistics, education and marital status were no longer significantly associated with Lifelines participation.

Model 3 included religious involvement. There is no evidence that attending religious services is associated with participation in Lifelines. Due to data restrictions, we could not include the religious denomination in the model for the whole period.⁵ As a robustness check, we added religious affiliation (Roman-Catholics, Protestants, Other, None) for the 2012-2017 period. This revealed that affiliation was not statistically significant ($p = 0.36$).

Models 4, 5 and 6 included the remaining variables as controls. We included subjective health status and a political interest scale and can observe that the impact of an individual's social capital remained basically the same. Finally, we included all controls in Model 6. It appears that subjective health status and political interest do not contribute separately to our understanding of why people become a Lifelines participant. This pattern did not change substantially if other controls are included in the models (not shown, see robustness checks).

5 In 2018, a redesign of the questions measuring religious involvement led to three alternative initial questions.

The impact of specific social capital indicators

As mentioned above, the social capital index was composed of 17 indicators. To determine which of these indicators were the main drivers of Lifelines participation, we employed additional models based on each indicator (see Table 3). In Model 7, we included the nine participation indicators. Subsequently, we added the eight trust indicators in Model 8. Finally, in Model 9, we included all participation and trust indicators. All of the models provide results with and without the inclusion of the controls. If applicable, the missing values were imputed by the median values (see Appendix 2 for an overview of the missing items). For voting, we assumed that people who were not eligible to vote did not vote.

Table 3. Logistic regression participation in Lifelines in the northern Netherlands, missing values imputed;

	Participation		Trust		Participation and trust		
	Model 7		Model 8		Model 9		
	Wald	Sig.	O.R.	Wald	Sig.	O.R.	
Contacts with family	4.8						5.1
Rarely or never (ref.)							
Less than once a month			0.74				0.72
At least once a month, but not weekly			0.84				0.83
At least once a week			1.08				1.08
Contacts with friends	0.8						0.7
Rarely or never (ref.)							
Less than once a month			1.3				1.28
At least once a month, but not weekly			1.12				1.12
At least once a week			1.13				1.12
Contacts with neighbors	8.7	*					9.1
Rarely or never (ref.)							*
Less than once a month		*	1.64				*
At least once a month, but not weekly		**	1.72				**
At least once a week		*	1.44				*

Informal help							
No (ref.)							
Yes	1.8		1.12		1.6		1.12
Volunteering							
No (ref.)							
Yes	3.0	†	1.17		2.6		1.16
Paid work							
No (ref.)							
Yes	3.5	†	1.25		3.6	†	1.25
Organizational activities							
No (ref.)							
Yes	13.3	***	1.37		12.6	***	1.36
Voting							
No (ref.)							
Yes	10.5	**	1.59		9.5	**	1.56
Political action							
No (ref.)							
Yes	0.6		0.93		0.5		0.94
Social trust							
No (ref.)							
Yes			3.9	*	1.21	1.8	1.14
Trust in army							
Not at all (ref.)			1.5			1.3	
Not so much					1.2		1.16
Quite a lot					1.26		1.23
A lot					1.17		1.15
Trust in police							
Not at all (ref.)			7.5	†		8.4	*
Not so much					1.07		1.09
Quite a lot					0.93		0.91
A lot					1.4		1.37
Trust in judges							
Not at all (ref.)			1.1			1.7	

Not so much										0.8		0.76
Quite a lot										0.79		0.74
A lot										0.82		0.77
Trust in civil servants				2.1							2.5	
Not at all (ref.)												
Not so much										1.02		1.04
Quite a lot										0.92		0.92
A lot										1.29		1.32
Trust in press				2.3							1.7	
Not at all (ref.)												
Not so much										1		0.99
Quite a lot										0.89		0.9
A lot										0.67		0.68
Trust in large companies				5.4							5.0	
Not at all (ref.)												
Not so much										0.9		0.89
Quite a lot										1.07		1.04
A lot										0.55		0.53
Trust in parliament				8.3	*						6.4	†
Not at all (ref.)												
Not so much					*					1.41		† 1.32
Quite a lot					†					1.37		1.27
A lot										0.66		0.62
Gender												
Male (ref.)												
Female	15.0	***	1.39	15.1	***	1.39				14.2	***	1.39
Age	86.1	***		123.5	***					85.9	***	
15-24 (ref.)												
25-34		***	2.65		***	2.77				***	2.68	
35-44		***	2.69		***	2.9				***	2.78	
45-54		***	4.03		***	4.24				***	4.2	
55-64		*	1.76		**	1.84				**	1.85	
65-74		*	1.62		†	1.55				*	1.7	
75+			0.71			0.63					0.72	
Marital status	8.0	*		9.4	*					8.2	*	

Married (ref.)									
Divorced			1.21		1.15			1.23	
Widowed			0.87		0.86			0.87	
Single		*	0.77		**	0.73		*	0.77
Education	3.9			6.7			4.0		
Elementary (ref.)									
Low			0.86		0.9				0.86
Middle			1.07		1.17				1.06
High			1.09		1.25				1.08
Highest			0.96		1.08				0.93
Standardised income household	18.2	***		27.7	***		17.9	***	
First (lowest) quartile (ref.)									
Second quartile			1.24		*	1.35			1.22
Third quartile		***	1.71		***	1.94		***	1.7
Fourth (highest) quartile		**	1.67		***	1.92		**	1.64
Migration background									
Dutch native (ref.)									
Migrant	2.7		0.74	4.6	*	0.67	2.2		0.76
Nagelkerke R square (%)	11.3			10.5			12.1		

Note: † $p < 0.10$ * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

The results in Model 7 show that participation in organizations and politics are the main drivers of the propensity to become a Lifelines participant, rather than social participation and providing informal help. To a lesser extent, volunteering work and having a paid job contributes to participation in Lifelines cohort studies ($p < 0.10$). In addition, having contacts with neighbours is relevant ($p < 0.05$). However, it should be noted that having no contacts at all was the only factor that had detrimental effects, which was a rather small group. This implies that connectedness to society in formal settings is more important for participation in a public good than informal networks involving family members, friends and others outside one's own household. With respect to trust, the results of Model 8 suggest that social trust and trust in parliament, and to a lesser extent trust in police, are the most important predictors of being a Lifelines participant, while trust in other institutions in the public and private domains had no impact.

We can also observe that, in this model, the controls point in the same direction, in comparison to those models in which the social capital composite index was included. In Model 9, all 17 trust and participation indicators were included, revealing that the results for participation changed little, with the exception of volunteering, which is no longer related to being a participant in the Lifelines cohort studies. In addition, social trust was no longer significantly associated with participating in Lifelines.

Robustness checks

We conducted several robustness checks to estimate the reliability of our findings. First, we controlled our findings by comparing the weighted data, with the aim of obtaining better representativeness of the Dutch population with the unweighted data. Second, we controlled for regional variation by including a variable that distinguished nine regions (three per province for Groningen, Friesland and Drenthe). The weighting, as well as the inclusion of the regional variable, had little effect on the odds ratios for the social capital and included controls. Third, our Lifelines sample might have underrepresented students, since participation required people to be settled in one place to some extent. However, our findings did not change after the exclusion of students from the S&W sample, which indicates that this is not a relevant bias in our selection. Finally, we replaced the social capital index based on five quintiles with one based on ten percentiles as well as adding an interval variable. All of the checks did not alter the findings or the conclusions related to the impact of social capital and other characteristics included in the models. We further explored whether the impact of social capital varied between subpopulations by including interactions in the models. No interactions appeared to be significant.

In addition, for Models 7, 8 and 9, we checked whether an alternative measurement of the participation and trust indicators would alter the results. Instead of the ordinal variables, we combined the answer options into binary variables to check whether the gain in degrees of freedom would lead to other findings. Social contacts were constructed as 'At least one contact per week' versus 'Less than one contact per week'. The seven institutional trust indicators were divided into 'No' (not so much, not at all) versus 'Yes' (quite a lot, a lot). If applicable, the missing values were imputed by the median value.

For participation, the findings were rather stable in comparison with Model 7, with the exception of social contacts. We found that regular contacts with neighbours was insignificant, while regular contact with family members was positively related to becoming a Lifelines participant ($p < 0.05$). Concerning the trust dimension (Model 8), the significance of trust in police and in parliament disappeared, and social trust was now related to Lifelines participation. The

findings based on the combined participation and trust indicators are also largely in line with the previous findings in Model 9. One exception is that volunteering was slightly related ($p < 0.10$) to being a Lifelines respondent. Furthermore, all relationships with trust were absent, including social trust and trust in police and parliament.

Conclusions and discussion

This paper examined the relationship between social capital and participation across public goods using unique data about social capital from Statistics Netherlands linked with the Lifelines biobank and longitudinal cohort study. We argued that participation in Lifelines can be understood as a particular form of collective action and prosocial behaviour. As such, being a respondent in Lifelines is a contribution to society and implies a sense of connectedness to other members of society. Thus, we assumed that people with more social capital would also be more likely to be participants in Lifelines.

Our examination showed the relevance of social capital to participation in cohort studies. We found that if a person was endowed with more social capital, such as greater levels of participation in formal associations and trust in others, the likelihood of them being a participant in Lifelines was higher. Our findings confirmed that social capital increases the probability that a respondent would participate in Lifelines. As this pattern was linear, this also implied that there is no 'dark side' of social capital (i.e. that very high levels of social capital might be detrimental to society; for an overview, see Portes, 2014), which is in line with other studies on the relationship between social capital and health (Wakefield & Poland, 2005; Gannon & Roberts, 2018). In addition, our results partly confirmed findings that high levels of societal trust and prosocial behaviour distinguished participants from those who decline to participate in research (Broekstra et al., 2021).

We discussed the social capital composite index, which was based on nine participation and eight trust indicators, and explored which indicators were most associated with becoming a Lifelines participant. Our analysis revealed that certain components matter more than others. In general, the social participation indicators seem to matter more than trust indicators. Only trust in police and in parliament had a significant, although rather weak, impact on participation in the Lifelines cohort study.

Yet these findings indicate that trust is more than just a prerequisite for participation in cohort studies: trust in authorities such as the police and parliament

is, at least to some extent, likely to foster contribution to public goods. Our interpretation of these findings is that trust in these political and other institutions may resemble the trust people have in the medical system and other established institutions, which was found to be an important aspect of being a member of the Lifelines cohort (Broekstra et al., 2021). Other studies showed that only composite scores on trust in society and biomedical research, for example, trust in hospitals, trust in government and trust in university researchers, explain the participation or non-participation of citizens in general in biobanks and in sharing medical data (Critchley et al., 2012; Gaskell et al., 2013; Kettis-Lindblad et al., 2006; Middleton et al., 2020; Milne et al., 2019). Thus, the inclusion of such specific trust indicators might provide further evidence of the relationship between trust in authorities and participation in the Lifelines cohort or similar scientific endeavours.

Furthermore, we found that formal participation in organizations and political participation mattered more than participating in informal activities with friends and neighbours and helping other people. More specifically, the fact that social contacts do not appear to be substantially relevant may indicate that being a Lifelines participant is not a frequent topic of discussion between people. Rather, what was most associated with Lifelines participation was a person's commitment to society, as expressed through being active in associations and politics. These results are in line with findings on formal donation behaviour as a predictor of participation, especially if individuals have a record of previous blood donation or registration for organ donation (Broekstra et al., 2021; Critchley et al., 2012; Gayet-Ageron et al., 2016). In contrast, the motive of helping friends or family by participating were low among Lifelines participants (Broekstra et al., 2021; Nobile et al., 2013).

Seen through the social capital lens, we also expected that membership of a religious community would increase the probability of being a Lifelines participant, given that social capital is more prevalent among religious groups. This expectation was not reflected in our results, as religious involvement had no unique impact on being a participant in the Lifelines studies. Although our findings do not support two recent studies showing a significant positive association with willingness to donate genomic and medical data, trust indicators were not taken into account as control variables in their analyses (Middleton, Milne, Thorogood et al., 2019; Voigt et al., 2020). However, our results do confirm those of another study that showed no independent effect of religiosity on participation in Lifelines, when taking into account multiple trust indicators (Broekstra et al., 2021). An explanation for our findings might be that the association of religiosity with participation is mediated by high levels of trust and being active in as-

sociations and in politics generally. We also found that the most religious group, those who attend religious services at least once a week, did not differ from the non-religious group with respect to participation in Lifelines. Our interpretation of this is that the rather small but very religious group in the Netherlands does not want to be involved in medical treatment and research in Lifelines. In a similar vein, this group more frequently refrains from organ donation than others (Schmeets & Peters, 2021).

Our paper also examined other variables related to becoming a Lifelines respondent. Women, people with a higher income and those aged between 25 and 64 were more willing to be Lifelines participants than men, the lower income groups and those aged between 1 and 24 and in the 65+ age group, respectively. The association of participation in biobanking with higher income has been previously reported, and it seems to be related to a reduced need for benefits (Ahram et al., 2014). Previous studies reported inconsistent findings regarding the positive effect of female gender and older age on participation in both biobanking (Broekstra et al., 2021) and surveys (Goyder, 1987; Groves & Couper, 1989; Stoop, 2005; Luiten, 2013). However, with regard to volunteering, it is known that age and gender differences in Europe are highly conditional, for example, they depend on the type of organization for which people volunteer (Gil-Lacruz et al., 2019). It might, therefore, be worthwhile to investigate the underlying mechanisms of the effects of age and gender, since mediating factors might explain the inconsistent findings, similar to religiosity.

Strengths and limitations

While our study made use of unique and high-quality data from the Dutch Survey on Social Cohesion and Wellbeing for the period 2012-2018 and from the Lifelines biobank and cohort study, there are several limitations to be taken into account. First, although the response rates for the survey are rather high – 63 to 70% in the various years – and weighting models were applied to reduce possible non-response bias, some bias remains. Recent studies showed that registered organ donors are overrepresented in the S&W sample and this is the same for Lifelines based on self-reported items about being organ and/or blood donor (Schmeets & Peters, 2020; Broekstra et al., 2021). Such a correlation reflects a prosociality bias among survey respondents, which reduces our ability to make inferences about the total population (Abraham, Presser & Helms, 2009).

The Lifelines participants in our sample were, moreover, involved in both the Lifelines cohort and the S&W survey, which might imply that they are extraordinarily motivated to contribute to research. This extraordinary motivation might

be associated with both social capital and the recruitment channel used to find participants for Lifelines. For example, participants who register themselves may have a higher level of social capital than participants who register through General Practitioners. Moreover, those participants registered through General Practitioners might have higher social capital levels than participants registered through family members. Unfortunately, we were not able to obtain the necessary data to distinguish between participants in these terms. This provides a further avenue for research on non-response bias in surveys and biobanking in the future. Furthermore, it would also be interesting to consider including organ donation in the weighting models; in our study, we were not permitted to use the organ donor register to assess non-response bias.

Second, our use of cross-sectional data limits our ability to address causality. The social capital index calculates the joint contribution of the 17 indicators to the first and second order constructs, and ultimately to the overall social capital score. The final limitation is that we do not have a specific indicator for trust in medical institutions, which may allow the precise determination of the level of trust within the relevant context of investigation. This would provide a further angle from which to explore the relationship between participation in public goods and other aspects of prosocial behaviour in the future.

References

- Abraham, K. G., Helms, S., & Presser, S. (2009). How social processes distort measurement: The impact of survey nonresponse on estimates of volunteer work in the United States. *American Journal of Sociology*, 114(4), 1129–1165.
- Ahram, M., Othman, A., Shahrouri, M., & Mustafa, E. (2014). Factors influencing public participation in biobanking. *European Journal of Human Genetics*, 22(4), 445–451. <https://doi.org/10.1038/ejhg.2013.174>
- Alesina, A., & La Ferrara, E. (2000). Participation in Heterogeneous Communities. *The Quarterly Journal of Economics*, 115(3), 847–904. <https://doi.org/10.1162/003355300554935>
- Bekkers, R. (2006). Traditional and Health-Related Philanthropy: The Role of Resources and personality. *Social Psychology Quarterly*, 69(4), p. 349–366.
- Bekkers, R.H.F.P. & Schuyt T.N.M. (2008). And who is your neighbor? Explaining denominational differences in charitable giving and volunteering in the Netherlands. *Review of Religious Research*, 50(1), 74–96.
- Bourdieu, P. (1983). The forms of capital. In J. Richardson (ed.), *Handbook of theory and research for the sociology of education*. New York: Greenwood Press.
- Brehm, O. (1993) *The Phantom Respondents. Opinion Surveys and Political Representation*. Michigan: Michigan University Press.
- Broekstra, R., Aris-Meijer, J., Maeckelberghe, E., Stolk, R., & Otten, S. (2021). Demographic and prosocial intrapersonal characteristics of biobank participants and refusers: the findings of a survey in the Netherlands. *European Journal of Human Genetics*. <https://doi.org/10.1038/s41431-020-0701-1>
- Caplan, A. L. (2016). Finding a solution to the organ shortage. *Cmaj*, 188(16), 1182–1183. <https://doi.org/10.1503/cmaj.151260>
- CBS, (2019a). *Open data over Sociale contacten* [social contacts] <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82249NED/table?ts=1601630728869>
- CBS, (2019b). *Open data over Donorregistratie* [organ donation rates] <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82249NED/table?ts=1601630384939>
- CBS, (2019c). *Open data over Vrijwilligerswerk* [volunteering in societal associations] <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82249NED/table?ts=1601630384939>
- CBS, (2019d). *Open data over Vertrouwen in mensen en organisaties* [Trust in people and organisations] <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82378NED/table?ts=1601630844260>
- CBS, (2021). <https://www.cbs.nl/nl-nl/cijfers/detail/80518ned?q=vertrouwen%20europees>
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, 95–120.

- Cote, S., & Healy, T. (2001). *The well-being of nations. The role of human and social capital*. Paris: Organisation for Economic Co-operation and Development.
- Critchley, C., Nicol, D., Otlowski, M., & Stranger, M. (2012). Predicting intention to biobank: A national survey. *European Journal of Public Health*, 22(1), 139–144. <https://doi.org/10.1093/eurpub/ckq136>
- De Hart, J. (2014). *Geloven binnen en buiten verband*. [Believe within and without context]. Den Haag: SCP.
- De Leeuw, E., Hox, J., & Luiten, A. (2018). International Nonresponse Trends across Countries and Years: An analysis of 36 years of Labour Force Survey data. *Survey Insights: Methods from the Field*, 1–11. <https://doi.org/10.11587/IZNRQ5>
- Flap, H. D., & Völker, B. (2004). Social networks and performance at work. In H. D. Flap & B. Völker (Eds.), *Creation and returns of social capital*, 172–196. London, UK: Routledge.
- Fukuyama, F. (2010) Het vertrouwen moet terug [Trust must return]. *Trouw*, 13-9-2010.
- Gannon, B. & Roberts, J. (2018). Social capital: exploring the theory and empirical divide. *Empirical Economics*. <https://doi.org/10.1007/s00181-018-1556-y>
- Gaskell, G., Gottweis, H., Starkbaum, J., Gerber, M. M., Broerse, J., Gottweis, U., et al. (2013). Publics and biobanks: Pan-European diversity and the challenge of responsible innovation. *European Journal of Human Genetics*, 21(1), 14–20. <https://doi.org/10.1038/ejhg.2012.104>
- Gayet-Ageron, A., Rudaz, S. & Perneger, T. (2016). Biobank attributes associated with higher patient participation: a randomized study. *European Journal of Human Genetics*, 25, 31–36. <https://doi.org/10.1038/ejhg.2016.132>
- Gil-Lacruz, M., Gracia-Perez, M. L., & Gil-Lacruz, A. I. (2019). Learning by Doing and Training Satisfaction: An Evaluation by Health Care Professionals. *Int. J. Environ. Res. Public Health*, 16(8), 1397. <https://doi.org/10.3390/ijerph16081397>
- Goyder, John (1987). *The silent minority : nonrespondents on sample surveys*, Cambridge polity press.
- Grootaert, C., & Narayan, D. (2004). Local institutions, poverty and household welfare in Bolivia. *World Development*, 32 (7), 1179–1198.
- Groves, R.M., & Couper, M.P. (1998). Nonresponse in household interview surveys. New York, NY: Wiley.
- Guillen, L., Coromina, L., & Saris, W. (2011). Measurement of social participation and its place in social capital theory. *Social Indicators Research*, 100, 331–350.
- Halpern, D. (2005). *Social capital*. Cambridge: Polity Press.
- Hanifan, L. J. (1920). *The Community Center*. Boston: Silver Burdett.
- Healy, K. (2004). Altruism as an organizational problem: The case of organ procurement. *American Sociological Review*, 69(3), 387–404.
- Healy, K. (2005). Do presumed-consent laws raise organ procurement rates. *De-Paul Law Review*, 55, 1017–1044.
- Hooghe, M., & Vanhoutte, B. (2011). Subjective well-being and social capital in

- Belgian communities. The impact of community characteristics on subjective well-being indicators in Belgium. *Social Indicators Research*, 100, 17–36.
- Hudson, J. (2006). Institutional trust and subjective well-being across the EU. *Kyklos*, 59(1), 43–62.
- Iisakka, L. (Ed.). (2006). *Social capital in Finland. Statistical review*. Helsinki: Statistics Finland.
- Iyver, S., Kitson, M., & Toh, B. (2005). Social capital, economic growth and regional development. *Regional Studies*, 39(8), 1015–1040.
- Keeley, B. (2007). *Human Capital. How what you know shapes your life*. Paris: OECD.
- Kettis-Lindblad Å., Ring L., Viberth E., & Hansson M. G. (2006). Genetic research and donation of tissue samples to biobanks. What do potential sample donors in the Swedish general public think? *Eur J Public Health*, 16, 433–440. <https://doi.org/10.1093/eurpub/cki198>
- Kim, H.-Y. (2018). Effects of social capital on collective action for community development. *Social Behavior and Personality: An international journal*, 46(6), 1011–1028.
- Klijs, B., Scholtens, S., Mandemakers, J. J., Snieder, H., Stolk, R. P., & Smidt, N. (2015). Representativeness of the LifeLines cohort study. *PLoS ONE*, 10(9), 1–12. <https://doi.org/10.1371/journal.pone.0137203>
- Knack, S. & Keefer, P. (1997) Does social capital have an economic pay-off? A cross country investigation. *Quarterly Journal of Economics*, 112 (4), 1251–1288.
- Kowalski, C. J. (1972). On the effects of non-normality on the distribution of the sample product-moment correlation coefficient. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 21(1), 1–12.
- Krokstad, S., Langhammer, A., Hveem, K., Holmen, T. L., Midthjell, K., Stene, T. R., et al. (2013). Cohort profile: The HUNT study, Norway. *International Journal of Epidemiology*, 42(4), 968–977. <https://doi.org/10.1093/ije/dys095>
- Ladin, K., Wang, R., Fleishman, A., Boger, M., & Rodrigue, J.R. (2015). Does Social Capital Explain Community-Level Differences in Organ Donation Designation? *The Milbank Quarterly*, 93(3), 609–641.
- Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. New York: Springer.
- Lochner, K., Kawachi, I., & Kennedy, B. P. (1999). Social capital: a guide to its measurement. *Health & Place*, 5(4), 259–270.
- Luiten, A., (2013), *Improving survey fieldwork with paradata*. PhD study, Universiteit Utrecht
- Luiten, A., Hox, J., & De Leeuw, E. (2020). Survey Nonresponse Trends and Fieldwork Effort in the 21st Century: Results of an International Study across Countries and Surveys. <https://doi.org/10.2478/jos-2020-0025>
- Matesanz, R., Domínguez-Gil, B., Coll, E., Mahillo, B., & Marazuela, R. (2017). How Spain Reached 40 Deceased Organ Donors per Million Population. *American Journal of Transplantation*, 17(6), 1447–1454. <https://doi.org/10.1111/>

- Meurs, P. (2008). *Sociaal vertrouwen: een kwestie van durf* [Social trust: a matter of courage]. In: SER/Verwey Jonker Instituut, Sociaal Vertrouwen. Eerste Verwey-Jonker/SER lezing [Social Trust. First Verwey-Jonker/SER lecture], 3 Sept 2008, 9–24.
- Merz, E. M., Van den Hurk, K., & De Kort, W. L. (2017). Organ Donation Registration and Decision-Making Among Current Blood Donors in the Netherlands. *Progress in Transplantation*, 27(3), 266–272.
- Middleton A, Milne R, Thorogood A, Kleiderman E, Niemiec E, Prainsack B, et al. (2019). Attitudes of publics who are unwilling to donate DNA data for research. *Eur J Med Genet.*;62:316–23. <https://doi.org/10.1016/j.ejmg.2018.11.014>
- Milne, R., Morley, K. I., Howard, H., Niemiec, E., Nicol, D., Critchley, C., et al. (2019). Trust in genomic data sharing among members of the general public in the UK, USA, Canada and Australia. *Human Genetics*, 138(11–12), 1237–1246. <https://doi.org/10.1007/s00439-019-02062-0>
- Neira, I., Vazquez, E., & Portela, M. (2009). An empirical analysis of social capital and economic growth in Europe (1980–2000). *Social Indicators Research*, 92, 111–129.
- Nieminen, T., Martelin, T., Koskinen, S., Simpura, J., Alanen, E., Härkänen, T., et al. (2008). Measurement and socio-demographic variation of social capital in a large population-based survey. *Social Indicators Research*, 85, 405–423.
- Nobile, H., Bergmann, M. M., Moldenhauer, J., & Borry, P. (2016). Participants' accounts on their decision to join a cohort study with an attached biobank: A qualitative content analysis study within two German studies. *Journal of Empirical Research on Human Research Ethics*, 11(3), 237–249. <https://doi.org/10.1177/1556264616657463>
- Nurul Islam G.M., Yew, T.S., & Viswanatha, K.K. (2014). Poverty and livelihood impacts of community based fisheries management in Bangladesh, *Ocean & Coastal Management* 96, 123–129.
- Nurul Islam G.M., Yew, T.S., Mustapha, N., Abdullah, R., & Viswanatha, K.K. (2011). Social capital, community based management, and fishers' livelihood in Bangladesh, *Ocean & Coastal Management* 54, 173–180.
- Parks, C. D., Joireman, J., & Van Lange, P. A. M. (2013). Cooperation, Trust, and Antagonism. *Psychological Science in the Public Interest*, 14(3), 119–165. <https://doi.org/10.1177/1529100612474436>
- Paxton, P. (1999). Is social capital declining in the United States? A multiple indicator assessment. *American Journal of Sociology*, 105(1), 88–127.
- Piersma, T.W., Bekkers, R., De Kort, W. & Merz, E.-M. (2019). Blood Donation across the Life Course: The Influence of Life Events on Donor Lapse, *Journal of Health and Social Behavior*, 60(2) 257–272.
- Ponthieux, S. (2004). The concept of social capital : a critical review 10th ACN Conference - Paris, 21–23 January 2004.
- Poortinga, W. (2006). Social relations or social capital? Individual and

- community health effects of bonding social capital. *Social Science & Medicine*, 63(1), 255-270.
- Portela, M., Neira, I., & Del Mar Salinas-Jiménez, M. (2013). Social capital and Subjective Wellbeing in Europe: A New Approach on Social Capital, *Social Indicators Research*, 114, 493-511.
- Portes, A. (1998). Social capital. Its origins and applications in modern sociology. *Annual Review of Sociology*, 24, 1-24.
- Portes, A. (2014). Downsides of social capital. *Proceedings of the National Academy of Sciences*, 111(52), 18407-18408.
- Putnam, R.D., Leonardi, R., & Nanetti, R. Y. (1993). *Making democracy work: civic traditions in modern Italy*. Princeton: Princeton University Press.
- Putnam, R. D. (1995a). Turning in, turning out: The strange disappearance of social capital in America. *Political Science and Politics*, 28, 664-683.
- Putnam, R. D. (1995b). America's declining social capital. *Journal of democracy* 6(1), 65-78.
- Putnam, R. (2000). *Bowling alone. The collapse and revival of American community*. New York: Simon & Schuster.
- Putnam, R.D., & Campbell D.E. (2012). *American Grace. How Religion Divides and Unites Us*. New York: Simon & Schuster.
- Rocco, I., Fumagalli, E., & Suhrcke, M. (2014). From Social Capital to Health – and Back. *Health economics*. <https://doi.org/10.1002/hec.2934>
- Schmeets, H., & Peters, F. W. C. (2016). De rol van religie in orgaandonatie. [The role of religion in organ donation] *Tijdschrift voor Recht, Religie en Beleid*, 7(3), 24-35.
- Schmeets, H. (2019). Religie en sociale cohesie [Religion and social cohesion], *Statistische Trends*, 1-21.
- Schmeets, H., & Te Riele, S. (2014). Declining Social Cohesion in The Netherlands? *Social Indicators Research*, 115(2), 791-812.
- Schmeets, H., & Peters, F (2021). The Impact of Social Capital on Organ Donation: Evidence from the Netherlands *Social Indicators Research*. <https://doi.org/10.1007/s11205-021-02666-y>.
- Scholtens, S., Smidt, N., Swertz, M. A., Bakker, S. J. L. L., Dotinga, A., Vonk, J. M., et al. (2015). Cohort Profile: LifeLines, a three-generation cohort study and biobank. *International Journal of Epidemiology*, 44(4), 1172-1180. <https://doi.org/10.1093/ije/dyu229>
- Sharp, C., & Randhawa, G. (2012). Social capital and Deceased Organ Donation, In: Randhawa (ed.) *Organ Donation and Transplantation*, London: IntechOpen, 115-136.
- Smidt, C.E. (2003). *Religion as social capital: producing the common good*. Waco: Baylor University Press.
- Stiglitz, J.E., Sen, A., & Fitoussi, J. (2009). Report by the Commissions on the Measurement of Economic Performance and Social Progress. (www.stiglitz-senfitoussi.fr)

- Stoop, I. 2005. *The Hunt for the Last Respondent*. The Hague: Social and Cultural Planning Office.
- Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., et al. (2015). UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Medicine*, 12(3), 1–10. <https://doi.org/10.1371/journal.pmed.1001779>
- Tenenhaus, M. (2004). PLS regression and PLS path modeling for multiple table analysis. COMPSTAT' 2004 Symposium, 1–12.
- UK Biobank Coordinating Centre. (2007). UK Biobank: Protocol for a large-scale prospective epidemiological resource UK Biobank Coordinating Centre Stockport. *Design*, 06(March), 1–112. <https://doi.org/10.1126/science.311.5767.1535c>
- Van Beuningen, J., & Schmeets, H. (2013). Developing a Social Capital Index for the Netherlands, *Social Indicators Research*, 113(3), 859–886.
- Van Beuningen, J., & Schmeets, H. (2015). De meetlat sociaal kapitaal. [The social capital measurement index] In: Schmeets, H. (red.) *Sociale samenhang: wat ons bindt en verdeelt* [Social cohesion: what us connects and divides]. Den Haag/Heerlen/Bonaire: CBS, 15–27.
- Van den Broek-Honingh, N., & De Jonge, J. (2018). Vertrouwen in de wetenschap – Monitor 2018. Den Haag: Rathenau Instituut. Verkregen van <https://www.rathenau.nl/nl/wetenschap-cijfers/kennisbenutting-en-impact/vertrouwen-de-wetenschap>
- Van Staa, T. P., Goldacre, B., Buchan, I., & Smeeth, L. (2016). Big health data: the need to earn public trust. *British Medical Journal*, 3636(July), i3636. <https://doi.org/10.1136/bmj.i3636>
- Voigt, T. H., Holtz, V., Niemiec, E., Howard, H. C., Middleton, A., & Prainsack, B. (2020). Willingness to donate genomic and other medical data: results from Germany. *European Journal of Human Genetics*, 1000–1009. <https://doi.org/10.1038/s41431-020-0611-2>
- Wakefield, S.E.L., & Poland, B. (2005). Friend, Family or Foe? Critical Reflections on the Relevance and Role of Social Capital in Health Promotion and Community Development. *Social Science & Medicine*, 60, 2819–2832. <https://doi.org/10.1016/j.socscimed.2004.11.012>
- Wetzels, M., Odekerken-Schröder, G., & Van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS Quarterly*, 33(March), 177–195.
- Wuthnow, R. (2002). Religious involvement and status-bridging social capital. *Journal for the Scientific Study of Religion*, 41(4), 669–684.

Appendix 1.: Measures for the 17 indicators in the social capital composite index

Social participation

1. Family contacts (Rarely or never (1) — at least once a day (5))
How often do you have contact with family members outside the home?
2. Contacts with friends (Rarely or never (1) — at least once a day (5))
How often do you have contact with friends or close acquaintances?
3. Contacts with neighbors (Rarely or never (1) — at least once a day (5))
How often do you have contact with neighbors?
4. Informal help (0 = No, 1 = Yes)
Have you provided any informal help in the last 4 weeks?

Organizational participation

5. Organizational activities (0 = No, 1 = Yes)
Do you participate in organizational activities at least once a month?
6. Work (0 = Less than twelve hours, 1 = 12 h or more)
How many hours per week do you spend on a paid job?
7. Volunteering (0 = No, 1 = Yes)
Have you done any volunteer work in the last 12 months for any of the following types of organizations:
 - Youth work
 - School
 - Health care organization
 - Sports organization
 - Hobby organization
 - Cultural organization
 - Church or mosque
 - Union or employee organization

- Political party or protest organization
- Social care organization
- Housing association
- Neighborhood organization
- Other

Political participation

8. Political action (0 = No, 1 = Yes)
In the last 5 years, have you participated in political actions or influenced politicians/civil servants via:
- radio, tv, newspaper
 - political parties or organizations
 - government organized meeting
 - personal contact with politician or civil servant
 - protest group
 - public protest or demonstration
 - internet, e-mail, SMS
 - otherwise
 - none of these
9. Voting (0 = No, 1 = Yes)
Have you voted in the most recent national elections?

Social trust

10. Social trust (0 = One can't be too careful, 1 = Most people can be trusted)
Indicate which statement you agree with most: most people can be trusted; one can't be too careful.

Organizational trust

11. Trust in army (1 = Not at all, 4 = A lot)
Indicate to which extent you trust the army
12. Trust in judges (1 = Not at all, 4 = A lot)
Indicate to which extent you trust judges

13. Trust in civil servants (1 = Not at all, 4 = A lot)
Indicate to which extent you trust civil servants

14. Trust in press (1 = Not at all, 4 = A lot)
Indicate to which extent you trust the press

15. Trust in police (1 = Not at all, 4 = A lot)
Indicate to which extent you trust the police

16. Trust in large companies (1 = Not at all, 4 = A lot)
Indicate to which extent you trust the large companies

Political trust

17. Trust in parliament (1 = Not at all, 4 = A lot)
Indicate to which extent you trust the Dutch parliament

Appendix 2. Descriptive statistics indicators social capital index 2012/2018 in the northern Netherlands

	Total N	Mean	S.D.	Minimum	Maximum	Missings
Contacts with family	5,598	3.76	0.6	1	4	65
Contacts with friends	5,605	3.63	0.71	1	4	58
Contacts with neighbors	5,593	3.37	0.99	1	4	70
Informal help	5,663	0.34	0.47	0	1	0
Volunteering	5,663	0.55	0.5	0	1	0
Paid work	5,663	0.52	0.5	0	1	0
Organizational activities	5,356	0.45	0.5	0	1	307
Voting	5,101	0.84	0.36	0	1	562
Political action	5,663	0.43	0.49	0	1	0
Social trust	5,374	0.62	0.49	0	1	289
Trust in army	5,040	2.64	0.74	1	4	623
Trust in judges	5,141	2.77	0.76	1	4	522
Trust in press	5,191	2.19	0.69	1	4	472
Trust in police	5,311	2.74	0.71	1	4	352
Trust in parliament	5,148	2.17	0.75	1	4	515
Trust in civil servants	5,120	2.34	0.72	1	4	543
Trust in large companies	5,023	2.25	0.73	1	4	640

Not participating in Lifelines						
	N	Mean	S.D.	Minimum	Maximum	Missings
Contacts with family	4,873	3.75	0.61	1	4	60
Contacts with friends	4,880	3.63	0.72	1	4	53
Contacts with neighbors	4,867	3.36	1.01	1	4	66
Informal help	4,933	0.33	0.47	0	1	0
Volunteering	4,933	0.54	0.5	0	1	0
Paid work	4,933	0.5	0.5	0	1	0
Organizational activities	4,658	0.44	0.5	0	1	275
Voting	4,408	0.83	0.37	0	1	525
Political action	4,933	0.42	0.49	0	1	0
Social trust	4,668	0.61	0.49	0	1	265
Trust in army	4,384	2.63	0.75	1	4	549
Trust in judges	4,465	2.76	0.76	1	4	468
Trust in press	4,507	2.18	0.69	1	4	426
Trust in police	4,615	2.74	0.72	1	4	318
Trust in parliament	4,472	2.16	0.76	1	4	461
Trust in civil servants	4,453	2.33	0.72	1	4	480
Trust in large companies	4,356	2.25	0.74	1	4	577

	Participating in Lifelines					
	<i>N</i>	<i>Mean</i>	<i>S.D.</i>	<i>Minimum</i>	<i>Maximum</i>	<i>Missings</i>
Contacts with family	725	3.83	0.5	1	4	5
Contacts with friends	725	3.65	0.66	1	4	5
Contacts with neighbors	726	3.47	0.85	1	4	4
Informal help	730	0.39	0.49	0	1	0
Volunteering	730	0.64	0.48	0	1	0
Paid work	730	0.69	0.46	0	1	0
Organizational activities	698	0.52	0.5	0	1	32
Voting	693	0.9	0.29	0	1	37
Political action	730	0.47	0.5	0	1	0
Social trust	706	0.69	0.46	0	1	24
Trust in army	656	2.68	0.69	1	4	74
Trust in judges	676	2.84	0.73	1	4	54
Trust in press	684	2.19	0.65	1	4	46
Trust in police	696	2.78	0.68	1	4	34
Trust in parliament	676	2.21	0.69	1	4	54
Trust in civil servants	667	2.36	0.68	1	4	63
Trust in large companies	667	2.26	0.7	1	4	63

Chapter 7

General Discussion

In this dissertation, we studied (from an interdisciplinary perspective) the underlying mechanisms of participation/non-participation in large-scale, centralized data repositories for scientific research and specifically of participation/non-participation in biobanks and large cohort studies. We identified two key factors of participation/non-participation, values and trust, in the context of participation/non-participation in a population-based biobank in the Northern part of the Netherlands, named Lifelines. Overall, the findings in this dissertation signal that individuals' levels of both prosocial orientation and values (defined as general concerns for others, themselves, and a common environment) and societal trust (such as general trust in the government or other individuals) can, to some extent, predict participation in a population-based biobank.

More specifically, qualitative research findings, described in Chapter 2, suggest that the decision of both participants and non-participants regarding their contribution to the biobank was partly motivated by prosocial considerations. Whereas participants perceived benefits and limited risks, non-participants and former participants were concerned about individual or societal harm. Chapter 3 adds to these findings with more qualitative research results by showing that differences in levels of trust depend strongly on four individual beliefs. First, the conviction that a large-scale data repository benefits the public. Second, the belief in the organizational aspect of data collection and storage, for example, the safe collection of data, the scale of the data repository, the sound reputation or trustworthiness of the responsible organization and the motives, behaviours and competence demonstrated by employees. Third, the belief that the data or insights, extracted from the collected data, are sensitive depending on range and context of repository. Fourth, the belief in the application of informed consent to retain control over personal data. Data suggest that these beliefs may influence trust and ultimately the decision to participate or not and the decision to withdraw from participation.

Our quantitative studies support and extend the findings of our qualitative studies. Chapter 4 indicates that prosocial intrapersonal characteristics, such as indicators of prosocial orientation and values or societal trust, were stronger predictors of participation in biobanking than demographic characteristics. Specifically, health-related values, trust in management and handling of data by commercial researchers appeared to be key factors in distinguishing participants from non-participants. Even though studies showed that motives for research participation are proso-

cial, we did not find a stronger general concern for others among participants compared to non-participants when considering other values and trust levels. Nor did we find a higher trust in researchers among participants beyond the context of biobanking. Chapter 5 confirms the fluidity of motives and levels of trust in deciding on participation, as motives for participation and withdrawal were related to perceived benefits or harm to society on the one hand and to the individuals themselves on the other. Levels of societal trust and of trust in data management and handling related differently to societal harm motives for withdrawal. Yet, in addition to general high levels of trust, the drivers of participants' willingness to accept data linkages for medical research included, a strong motivation to realize individual and particularly societal benefits through participation and weak concerns about possible individual or societal harm. Lastly, Chapter 6 tells us that if a person shows greater levels of commitment to society and trust in society, also known as social capital, the likelihood of them participating in both Lifelines and other data repositories for research is higher. While this commitment to society was reflected only in formal prosocial behaviours, such as voting or membership of non-profit organizations, trust in society reflected only a general trust in unknown others. Hence, the findings described in this chapter further emphasize the importance of enacted appreciation of the public context and high levels of trust in society or others for participation in large-scale, centralized data repositories for research.

In summary, the findings of this dissertation point to a dilemma in contributing to society without derogating individual interest and concerns. While participation in research data repositories for medical research requires participants to have strong prosocial values and orientations and a general trust in society (i.e. government and governmental researchers), the decision to participate relates both to appreciation of and experiences in the specific context of data repository and research (i.e. health and healthcare). Individuals' values and levels of trust in the context of data repositories and research might act as a facilitator or as a barrier, especially for participation via data linkage. Whereas health-related values (referred to as healthspheric values) facilitated participation in biobanks, environmental values (referred to as biospheric values) acted as barriers for participation. In a similar vein, trust in the context, referred to as trust in the public sphere of data repositories (i.e. data management and handling by governmental healthcare institutes and researchers), facilitated participation. In contrast, higher trust in the private sphere of data repositories (i.e. data management and handling by commercial enterprises) indicated that these may be barriers for participation. When data linkage comes into play, flows of information and data reach beyond the

context of data repositories for health research and the appreciation of this change in context seemed to associate particularly with the appreciation of the private sphere of data repositories and research. Our findings that participation requires commitment to and trust in society on the one hand and commitment to and trust in the context on the other is broadly in line with the literature [1–5]. The research presented is the first to indicate how trust may affect the willingness to participate and, in particular, the readiness to withdraw from participation, depending on beliefs about society and on the appreciation of the context of data collection.

Mixed motives for participation/non-participation

This dissertation underlines the variety of societal and individual motives for participation/non-participation in data repositories and the mechanism of reciprocity underlying participation. Our results indicate the importance of a common interest in participation by showing that the decision of participants and non-participants to contribute or not contribute to a biobank was, to some extent, motivated by a desire to help others, particularly future generations [Chapter 2 and 5]. These results, therefore, support studies that indicate that participation in biobanks is perceived as a contribution to science or to society [6, 7]. Moreover, our findings are in line with previous studies that indicate that contributing personal data to biobanks or to similar data repositories relates to prosocial attitudes and prosocial behaviours in the context of healthcare, such as positive attitudes towards biomedical research and experience with donating medical data [1, 3, 8], donating blood [9, 10] or donating organs [10]. Furthermore, we found a strong positive association between Lifelines participation and an individual's commitment to society as expressed through volunteering and other active contributions to society, such as participation in non-profit organizations and voting [Chapter 6]. This indicates that the act of participation in data repositories for health research relates to prosocial attitudes and prosocial behaviours beyond the biomedical context [Chapter 6], which is consistent with studies on participation in smart energy systems [11] and, on volunteering [12], and charitable giving [13]. Yet, participation is not merely about helping others since our findings indicate that a lack of general prosocial orientation or prosocial motives does not necessarily imply non-participation [Chapter 5]. Health-related values were independently associated with participation in Lifelines, which indicates the importance of individual interest [Chapters 2, 3 and 4]. In addition, we found, that participants from groups who are typically less likely to participate, such as males or less well-ed-

ucated individuals, were more motivated by individual benefits. In contrast, participants from groups who are likely to participate, such as females and older individuals, focused more on societal benefits [Chapter 5]. These findings confirm findings from other studies about these mixed motives for participation in biobanking or body donation for research, which emphasize the role of individual benefits as a motive in deciding on participation in research [6, 7, 14, 15].

In addition to a common interest to do good and an individual interest in health and healthcare, two other factors relating to individual interest are likely to play an independent role in deciding on participation. Respondents in our interview study explained that their inexperience with biomedical research or the lack of research participation by other people in their social network contributed to their non-participation [Chapter 2]. These factors of individuals' research participation experience and experiences with research participation in their social network show how participation is a social construct, in particular for the three-generation population-based biobank under study.

Participation in research, like organ donation, can be perceived as a form of collective action [16–18]. This perception is, however, somewhat problematic as it leaves out the possibility that participation in research might lead to a long-term loss or harm. Participation is defined, from this perspective, as a cooperation among individuals to achieve a long-term gain in knowledge through contribution to research. Our results in Chapter 3 indicate, however, that some non-participants and former participants perceived participation in large-scale data repositories differently and interpreted the goal of gaining knowledge as a potential threat to individuals. Hence, their non-participation may also be perceived as a collective act aimed at preventing a gain in knowledge via research. This positive normative view on decisions to participate in biomedical research excludes a positive interpretation of non-participation while a definition could benefit from a more neutral and non-normative approach.

A slightly alternative approach is to define participation as a form of prosocial behaviour since it benefits other individuals or groups of individuals [19–21]. This approach might be more suitable since achievement of a common goal is not a condition. According to this definition, individuals support others by participating based on a motive to help others, which we indeed confirmed in our results [Chapter 2 and 5]. However, we also found that motives for individual benefits, just as trust and values, played an important role in the decision to participate [Chapters 2 and 5]. This seems to imply that participation is not a choice to act prosocially without taking self-interest into consideration. Moreover, it is not possible to define non-participation as prosocial, even though individuals assumed

that participation would pose a potential threat to individuals or groups of individuals.

Although both approaches to participation are imperfect by including normative assumptions about participation/non-participation, they both provide a framework for explaining how people might cope differently with the key social dilemma associated with participation: contributing to society without derogating individual interest and concerns. For the collective action approach, the common goal of creating or preventing the creation of new knowledge from data repositories through research might depend on individuals' views on societal and individual benefits and harms and their concerns about the context of data repositories. Conversely, in the prosocial behaviour approach, participation might conflict with common and individual interests or concerns, depending on the view on society, data repositories and research. For example, participation is viewed as a common aim or as helpful since the data repositories for research are believed to contribute to improving the future wellbeing of individuals. This might explain why trust, in addition to prosocial orientations and values, is key in distinguishing participants in data repositories and non-participants. Individuals must decide to participate in data repositories or not and rely on their levels of societal trust and trust in data management and handling to estimate the future consequences of the data repositories and research.

Trust as a predictor of participation/non-participation

Views on benefits and harms of participation and non-participation in a large-scale, centralized data repository were, according to our results, partly derived from individuals' appreciation of and experiences in society, such as healthspheric or biospheric values and societal trust. As a result, deciding on participation relates to issues of trust and experiences in society regardless of the level of commitment to society. According to a systematic review about attitudes towards the reuse of health data in the European Union, participation in data repositories is always conditional [22]. Our data suggest that even though trust in society and trust in the relevant context of data repositories are strongly correlated, both predict unique variance in participation/non-participation [Chapters 4 and 5]. Participants had more trust in government and other individuals [Chapter 4]. At the same time, people who generally trust other individuals are more willing to be a participant in multiple data repositories [Chapter 6]. We found that trust could simultaneously act as a facilitator and as a barrier for participation in biobanking, depending on the context, since non-participants in our study had less societal trust and yet more trust in

commercial researchers compared to participants [Chapter 4].

Our findings support a view of trust as a conditional and fluid concept [1, 23, 24]. This is in line with recent studies that show a high variability in trust in society, in particular in the healthcare context [3, 25]. Our results fit a definition of trust in society as a resource to cope with uncertainties in benefits and harms of data repositories and research [26, 27], for example about involvement of commercial enterprises in research. Differences between participants' and non-participants' beliefs about society, research and data repositories (e.g. trust in society, trust in research institutes or trust in commercial enterprises) have been identified in other studies as well [5, 8, 28, 29]. When applying data linkage or sensor technology, the differences between participants' and non-participants' levels of trust, such as trust in others, trust in government or trust in data management and handling in research, seemed to affect participation in data repositories even more. Participants with higher levels of trust in the entire society (i.e. in government, in commercial enterprises, in other individuals and in commercial researchers) indicated to be more willing to contribute data via data linkage or sensor technology than others, especially when these were combined with strong motives for participation and weak motives for withdrawal [Chapter 5].

Hence, to increase the probability that individuals will participate in large-scale, centralized data repositories that apply these technologies, high and broad levels of trust in society should be present as a resource for deciding in favour of participation. For example, conducting research with data repositories using sensor technology or data linkage requires trust in the public domain (i.e. government and governmental researchers) as well as in the private domain (i.e. commercial enterprises and commercial researchers). The application of these sensor technologies or data linkages can change the flow of information and, related to this, the processes of research or the ideas about the constitution of knowledge [30–32]. For example, it might change previously clear distinctions between research and other purposes or change the current roles of public enterprises and commercial enterprises in research. This change of information flow and related ideas about research and knowledge seems to create issues of trust in data repositories for research, especially for people who are currently convinced to participate in data repositories, as they do not trust the private domain. Their trust is likely to be based on a confidence in the current context of data repositories that is characterized by the public sphere, which might be associated with the underestimation of the current risks of data repositories and the existing supporting role of commercial enterprises in research [29, 33]. If the context of data repositories and research changes information flows and ideas about research

and knowledge, this might create cognitive dissonance between current beliefs about sources of harm and new realities that foster withdrawal or non-participation, in particular for those who were more open to change in the first place [34].

Implications for recruitment and retainment of participants

Altogether, our findings indicate how important positive beliefs about society and Even though individual benefits are commonly accepted as motives for participation in data repositories in other contexts, these motives are still quite often perceived with scepticism when it comes to participation in data repositories for scientific research. One of the reasons for this scepticism is that accepting incentives for individual benefits as a motive for participation would allow ethical harms, for example exacerbation of inequality in healthcare [35, 36]. Our results show, however, that values and individual benefit motives are part of the decision-making process about participation, which emphasizes the inevitable reciprocal character of participation [Chapters 2 and 5]. Other studies also showed that motives for participation tend to be a mix of common and individual interests and purposes [14, 37].

Therefore, this topic needs to be addressed and evaluated by data repositories and researchers. Our findings show that groups who are on average difficult to recruit, such as individuals with lower educational backgrounds or males, pursue individual benefits in particular through their participation [Chapter 5]. This currently creates an inequality in research participation, which in the first instance implies indirect harm of less knowledge about these groups of individuals in the long term. When focusing on the medical context, this inequity is, however, likely to become an inequity in future healthcare since health data repositories steadily evolve to simultaneously facilitate research and clinical care, e.g. in learning health systems. These learning health systems aim to continuously improve data repositories and tailor treatments to individuals [38]. As a result, the strict distinction between research and other purposes is getting blurred, which means that the reality of research on large-scale, centralized data repositories with therapeutic or diagnostic purposes might no longer be a misconception. Research participation may directly improve participants' health, for example via genetic screening for hereditary risk. Recently, a study in Estonia showed how participants of a biobank benefitted individually from genetic screening for hereditary breast cancer and ovarian cancer risks [39].

Medical scientific researchers are obliged to prevent participants from

having ‘misconceptions’ that their research has therapeutic or diagnostic objectives [40]. Yet, expectations that research provides individual satisfaction and might provide individual health benefits in the future are commonly accepted [7]. For example, some Lifelines participants and former participants acknowledged that a reason to participate was the potential return of alarming individual results to their general practitioner [Chapters 2 and 3]. This distinction between intolerable conception of therapeutic or diagnostic objectives and tolerable hope of individual health benefits might be perceived as strict and somewhat arbitrary, especially for large-scale, centralized data repositories. The benefits and objectives of data repositories continuously develop at a rapid pace during their time of existence. This strict simplification of the complexity of benefits and risks of large-scale, centralized data repositories and research for societies and individuals might obscure decision-making about participation in the long run since participation is based on individual beliefs and hopes rather than on facts or shared views and determined goals. It could subsequently undermine a key strength of large-scale, centralized data repositories in a health context, as participation enables individuals to take preventive measures as long as they participate. The groups who currently do not participate might be the ones who could benefit most from these preventive measures. Hence, the hybrid space of research and other purposes, e.g. therapeutic or diagnostic, urges us to rethink the positive role of individual benefits that relate to the improvement of health in return for research participation, in particular those that relate to biobanking [41].

It might, therefore, be acceptable to convince individuals to participate in data repositories by relating to motives other than helping others in society, for example, the motive to adhere to a norm in their social network or to fulfil individual interests. By emphasizing the strengths of the data repositories, both in terms of adding valuable information to information that is available in existing repositories and by providing clear individual benefits to the respondents (e.g., regular health check-ups), a wider variety of individuals could be recruited for participation [42]. This could perhaps lead to an increasing number of people who have experience with participating in data repositories for research and to more participants from different social networks, thereby further increasing participation in data repositories. This might also empower groups that are currently left out in research, because of a strict, yet somewhat arbitrary, interpretation of tolerable individual benefits of participation [43–45].

Implications for governance of data repositories

Altogether, our findings indicate how important positive beliefs about society and the trustworthiness of the research context are in deciding about participation in data repositories, especially when data optimization or individual benefit via data linkage is offered. Whereas some groups within the population may be easier to recruit, other groups may be easier to retain when these contextual changes in biobanking occur. This supports the need for tailored recruitment and retainment strategies in an evidently changing context of data repositories [46]. Nowadays biobanks and other data repositories cross borders and contexts via complex systems, especially when applying data linkage and sensor data technologies. Researchers often rely on informed consent forms to make sure that participants are well-informed about the goals and potential risks of participation. However, studies showed that informed consent is often poorly understood [47] and that knowledge about technology seems to play a limited role in the acceptance of technology [48]. Participation in research endeavours is fostered mainly by a societal trust that, in our studies, does not include the private sphere, while awareness about current and potential risks of data sharing in biobanking and other large-scale research is still limited [3, 5, 8, 49]. Data should be collected in a way that enables participants to have sufficient trust in data repositories while acknowledging the potential societal and individual consequences of participation, for example about genetic research or involvement of commercial enterprises. Participants should not be lured into being overconfident that potential individual consequences do not exist or matter [33]. Otherwise, this trust in research is based on false premises and misconceptions, which might break down this trust in the long term [50].

The focus for data repositories in scientific research should therefore shift to systems of recruitment for and governance of data that proactively enable informed decision-making about participation. More transparency and proactive communication strategies should be employed by organizations that are responsible for data repositories in order to be worthy of their trust. This goes further than merely confirming that participants have sufficient knowledge about certain risks and benefits in the respective research via an informed consent form at the start of the participation or about acting on incidents such as data leakage. Organizations that are responsible for data repositories should ensure that governance systems and consequences of participation/non-participation in research are clearly and continuously communicated within organizations and with participants and others in society, especially in population-based data repository. Information officers, data stewards or data custodians from those organisations might play a key role in such new approach [51, 52].

This relationship-based approach could help to correct and prevent misconceptions about research and turn individuals' given trust into a rich resource to cope with uncertainties about data reposition and research [26], for example about the involvement of commercial enterprises in research and the related increasing concerns about biosamples or data becoming commercial products or about privacy violations [29, 33].

The responsibility to ensure that decisions on participation are informed and based on facts does not only lie in the hands of the organizations that set up data repositories. Governments and public institutes that represent research on a national and/or international level are key to the robust governance of research data repositories. They could and should help with informing, discussing and deciding about responsibilities and limitations of risks of potential societal and individual harm of data repositories. Only then can informed consent at the institutional level imply someone being informed about what they consent to in research participation but also being informed about when and how that individual will be notified of situations that might require a renewal or adaptation of the consent given. Moreover, governments and public institutes that represent research should develop adaptive and sustainable systems for protecting research participants and non-participants from harm concurrently with large-scale, centralized data repositories, for example through harm mitigation bodies or legislation that prevents insurance exclusion [53]. When developing these systems, the national and cultural context must be taken into account, such as levels and mechanisms of trust, general values, and distinctions between the private and public domains [29, 54].

A stepping stone could be increasing actual or potential participants' awareness of data impact by applying more visualizations of data processes, information flows and potential consequences by researchers themselves. Current information and communication systems provide these opportunities. National and international governing bodies, such as data protection authorities or ethical review boards, should help to further clarify opportunities and limitations for data repositories or research and strengthen proactive supervision of data repositories and their further use. Ensuring a better overview of the available data and databases might help to provide more guidance and clarity to the public and might foster the efficient use of data repositories. The establishment of data principles such as *FAIR* (*Findable, Accessible, Interoperable, and Reusable*) [55], or research infrastructure initiatives such as the Dutch *HEALTH-Research Infrastructure*, could be important stepping stones towards an improvement of insight into and consensus on opportunities and limitations of data repositories.

Limitations

Our studies yielded insights into which factors affect individuals' willingness to contribute their personal data to large-scale, centralized data repositories for scientific research and in particular to biobanks, although the studies had some limitations. First, all respondents were participants in a Dutch large-scale, centralized data repository, which is a population-based biobank. Therefore, our findings may be specific to the Dutch societal context of data repositories, especially in the context of biomedical research. For example, distinctions between the public and private sectors may vary in other countries and contexts [24], as well as concerns about privacy [25]. Hence, the mechanisms underlying participation in data repositories, such as trust and values, could be different in other countries, as shown recently in a global study about genomic data sharing [5]. The Dutch are known to have a high engagement and trust in biobanking and fewer data privacy concerns compared to inhabitants of other European countries [1]. Moreover, they tend to perceive health and medical care as a priority for science, research and innovation [56]. These attitudes, which are characteristic of Dutch people, may also apply to samples from different types of biobanks and cohort studies. Second, causal effects are difficult to capture through individual in-depth interviews and exploratory cross-sectional surveys. This issue could be resolved by using more experimental or longitudinal designs in future quantitative studies, together with qualitative confirmatory investigations of these processes. Third, although our findings were statistically significant, the explained variance of our models was limited. Moreover, not all our measurements relied on validated scales, which further limits the generalization and validity. Therefore, while our findings provide an insight into the factors that may affect individuals' willingness to contribute their personal data to large-scale, centralized data repositories for scientific research, and specifically to biobanks, these factors require and deserve further investigation. Nevertheless, the current findings may already provide valuable insights for the development of effective recruitment and retainment strategies and for the governance of large-scale centralized data repositories.

Conclusion

Values and trust provide rich grounds for understanding decisions regarding participation or non-participation, or withdrawal from data repositories for research. If participation decisions are made with expectations that are in line with reality and within thorough governance struc-

tures, these decisions honour a broad trust in society. It is not so much about the fact *that* a request for data is made, but more about *how* this request is made and *what* the conditions surrounding the data request are. We hope that our findings will provide a relevant stepping stone for practical interventions directed at properly addressing potential participants' ethical and practical concerns and, subsequently, at enhancing and maintaining their trust and the probability that they will provide their personal data for scientific research. In this way, our present research, together with follow-up studies in this field, may contribute to the further spread of rich, large-scale data repositories on health-related and other societally relevant issues.

References

- [1] Gaskell G, Gottweis H, Starkbaum J, et al. Publics and biobanks: Pan-European diversity and the challenge of responsible innovation. *Eur J Hum Genet* 2013; 21: 14–20.
- [2] Middleton A, Milne R, Howard H, et al. Members of the public in the USA, UK, Canada and Australia expressing genetic exceptionalism say they are more willing to donate genomic data. *Eur J Hum Genet* 2019; 424–434.
- [3] Milne R, Morley KI, Howard H, et al. Trust in genomic data sharing among members of the general public in the UK, USA, Canada and Australia. *Hum Genet* 2019; 138: 1237–1246.
- [4] Steinsbekk KS, Ursin LØ, Skolbekken JA, et al. We're not in it for the money-lay people's moral intuitions on commercial use of 'their' biobank. *Med Heal Care Philos* 2013; 16: 151–162.
- [5] Middleton A, Milne R, Almarri MA, et al. Global Public Perceptions of Genomic Data Sharing : What Shapes the Willingness to Donate DNA and Health Data ? *Am J Hum Genet* 2020; 1–10.
- [6] Locock L, Boylan AMR. Biosamples as gifts? How participants in biobanking projects talk about donation. *Heal Expect* 2016; 19: 805–816.
- [7] Nobile H, Vermeulen E, Thys K, et al. Why do participants enroll in population biobank studies? A systematic literature review. *Expert Rev Mol Diagn* 2013; 13: 35–47.
- [8] Critchley C, Nicol D, Otlowski M, et al. Predicting intention to biobank: A national survey. *Eur J Public Health* 2012; 22: 139–144.

- [9] Raivola V, Snell K, Helén I, et al. Attitudes of blood donors to their sample and data donation for biobanking. *Eur J Hum Genet* 2019; 27: 1659–1667.
- [10] Gayet-Ageron A, Rudaz S, Perneger T. Biobank attributes associated with higher patient participation: A randomized study. *Eur J Hum Genet* 2016; 25: 31–36.
- [11] van der Werff E, Steg L. The psychology of participation and interest in smart energy systems: Comparing the value-belief-norm theory and the value-identity-personal norm model. *Energy Res Soc Sci* 2016; 22: 107–114.
- [12] Hyde MK, Knowles SR. What predicts Australian university students' intentions to volunteer their time for community service? *Aust J Psychol* 2013; 65: 135–145.
- [13] Van Lange PAM, Bekkers R, Schuyt TNM, et al. From Games to Giving: Social Value Orientation Predicts Donations to Noble Causes. *Basic Appl Soc Psych* 2007; 29: 375–384.
- [14] Forsberg JS, Hansson MG, Eriksson S. Changing perspectives in biobank research: From individual rights to concerns about public health regarding the return of results. *Eur J Hum Genet* 2009; 17: 1544–1549.
- [15] Bolt S, Eisinga R, Venbrux E, et al. Personality and motivation for body donation. *Ann Anat* 2011; 193: 112–117.
- [16] Healy K. Altruism as an Organizational Problem: The Case of Organ Procurement. *Am Sociol Rev* 2004; 69: 387–404.
- [17] Healy K. Do presumed-consent laws raise organ procurement rates. *DePaul Law Rev* 2005; 55: 1017.
- [18] Ladin K, Wang R, Fleishman A, et al. Does Social Capital Explain Community-Level Differences in Organ Donor Designation? *Milbank Q* 2015; 93: 609–641.
- [19] Bekkers R. Traditional and health-related philanthropy: The role of resources and personality. *Soc Psychol Q* 2006; 69: 349–366.
- [20] Merz EM, Van Den Hurk K, De Kort WLAM. Organ donation registration and decision-making among current blood donors in the Netherlands. *Prog Transplant* 2017; 27: 266–272.
- [21] Piersma TW, Bekkers R, de Kort W, et al. Blood Donation across the Life Course: The Influence of Life Events on Donor Lapse. *J Health Soc Behav*

2019; 60: 257–272.

- [22] Skovgaard LL, Wadmann S, Hoeyer K. A review of attitudes towards the reuse of health data among people in the European Union: The primacy of purpose and the common good. *Health Policy (New York)* 2019; 123: 564–571.
- [23] Broekstra R, Aris-Meijer J, Maeckelberge E, et al. Trust in Centralised Large-Scale Data Repository: A Qualitative Analysis. *J Empir Res Hum Res Ethics*.
- [24] Sheikh Z, Hoeyer K. “That is why I have trust”: unpacking what ‘trust’ means to participants in international genetic research in Pakistan and Denmark. *Med Heal Care Philos* 2018; 21: 169–179.
- [25] Voigt TH, Holtz V, Niemiec E, et al. Willingness to donate genomic and other medical data: results from Germany. *Eur J Hum Genet* 2020; 28: 1000–1009.
- [26] Luhmann N. *Trust and Power*. John Wiley & Sons Inc., <https://books.google.nl/books?id=WZu4AAAAIAAJ> (1979).
- [27] Mújdricza F. The roots of trust. *Eur J Ment Heal* 2019; 14: 109–142.
- [28] Middleton A, Milne R, Thorogood A, et al. Attitudes of publics who are unwilling to donate DNA data for research. *Eur J Med Genet* 2019; 62: 316–323.
- [29] Richter G, Borzikowsky C, Lesch W, et al. Secondary research use of personal medical data: attitudes from patient and population surveys in The Netherlands and Germany. *Eur J Hum Genet*. Epub ahead of print 1 October 2020. DOI: 10.1038/s41431-020-00735-3.
- [30] Boyd D, Crawford K. Critical Questions for Big Data. *Information, Commun Soc* 2012; 15: 662–679.
- [31] Barocas S, Nissenbaum H. Big Data ’ s End Run around Anonymity and Consent. In: Julia L, Stodden V, Stefan B, et al. (eds) *Privacy, Big Data and the Public Good*. Cambridge; New York: Cambridge University Press, pp. 44–75.
- [32] Mayer-Schönberger V, Cukier K. *Big data : a revolution that will transform how we live, work and think*. London SE: John Murray, http://bvbr.bib-bvb.de/8991/F?func=serviceetdoc.library=BVB01etdoc.number=025907878etline_number=0001etfunc_code=DB_RECORDSetservice_type=MEDIA (2013).
- [33] Ursin L, Ytterhus B, Christensen E, et al. «If you give them your little finger, they’ll tear off your entire arm»: losing trust in biobank research. *Med Heal Care Philos* 2020; 23: 565–576.
- [34] Richter G, Borzikowsky C, Lieb W, et al. Patient views on research use

of clinical data without consent: Legal, but also acceptable? *Eur J Hum Genet* 2019; 27: 841–847.

- [35] Sass RG. Toward a More Stable Blood Supply: Charitable Incentives , Donation Rates , and the Experience of September 11. *Am J Bioeth* 2013; 13: 38–45.
- [36] Martin D. Charitable Incentives for Blood Donation Are Promising, But Require Careful Consideration. *Am J Bioeth* 2013; 13: 52–54.
- [37] Nobile H, Bergmann MM, Moldenhauer J, et al. Participants' accounts on their decision to join a cohort study with an attached biobank: A qualitative content analysis study within two German studies. *J Empir Res Hum Res Ethics* 2016; 11: 237–249.
- [38] Rivera SM, Brothers KB, Cadigan RJ, et al. Modernizing Research Regulations Is Not Enough: It's Time to Think Outside the Regulatory Box. *Am J Bioeth* 2017; 17: 1–3.
- [39] Leitsalu L, Palover M, Sikka TT, et al. Genotype-first approach to the detection of hereditary breast and ovarian cancer risk, and effects of risk disclosure to biobank participants. *Eur J Hum Genet*. Epub ahead of print 23 November 2020. DOI: 10.1038/s41431-020-00760-2.
- [40] Horng S, Grady C. Misunderstanding in Clinical Research: Distinguishing Therapeutic Misconception, Therapeutic Misestimation, and Therapeutic Optimism. *IRB Ethics Hum Res* 2003; 25: 11–16.
- [41] Johnsson L, Helgesson G, Hansson MG, et al. Adequate trust avails, mistaken trust matters: On the moral responsibility of doctors as proxies for patients' trust in biobank research. *Bioethics* 2013; 27: 485–492.
- [42] Kaufman D, Murphy J, Scott J, et al. Subjects matter: A survey of public opinions about a large genetic cohort study. *Genet Med* 2008; 10: 831–839.
- [43] Bekkers R, Wiepking P. *A literature review of empirical studies of philanthropy: Eight mechanisms that drive charitable giving*. 2011. Epub ahead of print 2011. DOI: 10.1177/0899764010380927.
- [44] Ridgeway JL, Han LC, Olson JE, et al. Potential Bias in the Bank: What Distinguishes Refusers, Nonresponders and Participants in a Clinic-Based Biobank? *Public Health Genomics* 2013; 16: 118–126.
- [45] Cheung KL, Ten Klooster PM, Smit C, et al. The impact of non-response bias due to sampling in public health studies: A comparison of voluntary versus mandatory recruitment in a Dutch national survey on adolescent health. *BMC Public Health* 2017; 17: 1–10.
- [46] Sanderson SC, Brothers KB, Mercaldo ND, et al. Public Attitudes toward Consent and Data Sharing in Biobank Research: A Large Multi-site Experimental Survey in the US. *Am J Hum Genet* 2017; 100: 414–427.

- [47] Broekstra R, Maeckelberghe ELM, Stolk RP. Written informed consent in health research is outdated. *Eur J Public Health* 2017; 27: 194–195.
- [48] Connor M, Siegrist M. Factors influencing people’s acceptance of gene technology: The role of knowledge, health expectations, naturalness, and social trust. *Sci Commun* 2010; 32: 514–538.
- [49] Bossert S, Kahrass H, Strech D. The public’s awareness of and attitude toward research biobanks – A regional German survey. *Front Genet* 2018; 9: 1–11.
- [50] O’Neill O. Linking Trust to Trustworthiness. *Int J Philos Stud* 2018; 26: 293–300.
- [51] Allen J, Adams C, Flack F. The role of data custodians in establishing and maintaining social licence for health research. *Bioethics* 2019; 404–409.
- [52] Kowalczyk S, Shankar K. Data sharing in the sciences. *Annu Rev Inf Sci Technol* 2011; 45: 247–294.
- [53] McMahon A, Buyx A, Prainsack B. Big Data Governance Needs More Collective Responsibility: The Role of Harm Mitigation in the Governance of Data Use in Medicine and Beyond. *Med Law Rev* 2020; 28: 155–182.
- [54] Gaskell G, Stares S, Allansdottir A, et al. *Europeans and Biotechnology in 2010 – Winds of change?* 2010.
- [55] Wilkinson MD, Dumontier M, Aalbersberg IJJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 2016; 3: 160018.
- [56] Eurobarometer. *Public perceptions of science, research and innovation*. Epub ahead of print 2014. DOI: 10.2777/95599

Algemene samenvatting

Waarden van vertrouwen en deelname in wetenschappelijke gegevensverzamelingen.

Gedurende de laatste dertig jaar is het opslaan van gegevens veranderd van fysiek naar digitaal. Deze digitalisering heeft het genereren, opslaan, gebruik en de uitwisseling van gegevens gestimuleerd. Hierdoor zijn digitale gegevensverzamelingen een cruciale bron voor wetenschappelijk onderzoek geworden. Van- daag de dag kan een als maar groeiende berg aan persoonlijke gegevens mak- kelijker verkregen worden via sensor technologieën, gegevensuitwisselingen, of open bronnen zoals sociale media. Bovendien produceert een groeiend aantal mensen en apparaten digitale gegevens. Deze digitale gegevensverzamelingen brengen zo kansen voor het doen van een nieuwe soort onderzoek, die verder gaan dan onderzoek met niet-digitale gegevensverzamelingen. Dit soort onder- zoek kijkt namelijk naar verbanden in grote gegevensverzamelingen die bestaan uit verschillende type gegevensverzamelingen en soorten onderzoek.

Biomedisch onderzoek is één van de onderzoeksvelden die geraakt wordt door deze nieuwe onderzoeksmogelijkheden, omdat het inzetten van big data doorslaggevend kan zijn voor het begrijpen en verbeteren van gezondheid en ge- zondheid gerelateerd gedrag van een individu. Hiervoor is het wel nodig dat veel mensen een grote diversiteit aan gegevens bijdragen. Juist daardoor worden digitale gegevensverzamelingen voor onderzoek grootschalig. Bovendien geven deze gegevensverzamelingen op één plek toegang tot verschillende soorten ge- gevens in hun verzameling. Goede voorbeelden van deze grootschalige, gecen- traliseerde gegevensverzamelingen in biomedische onderzoek zijn biobanken en grote cohort studies. Deze banken bewaren grote hoeveelheden monsters van biologisch materiaal gedurende een langere periode. Daarnaast bewaren zij bijbehorende gegevens, zoals ingevulde vragenlijsten en verschillende metin- gen. Zodoende bevatten deze banken voor een langere periode veel gevoelige gegevens, bijvoorbeeld gezondheidsgegevens of genetische gegevens, die een gevolg kunnen hebben voor iemand zijn levensomstandigheden, kwaliteit van leven en mogelijkheden in de maatschappij.

Veel mensen zijn het er over eens dat deze verandering in grootschalige en gecentraliseerde gegevensverzamelingen voor onderzoek ontwrichtend is, om- dat dit de stroom van informatie – zeker gevoelige informatie – en ideeën over onderzoek doen en het opdoen van kennis verandert. De kennis en informatie

die opgedaan wordt met deze gegevensverzamelingen kan nieuwe voordelen bieden voor individuen en samenlevingen. Tegelijkertijd kunnen deze mogelijkheden gevaren met zich meebrengen voor diezelfde individuen en samenlevingen. Zo kan het begrijpen van gezondheid en gezondheidsgedrag van individuen over een langere periode verstreckende gevolgen hebben, niet alleen voor de klinische praktijk, maar ook voor fundamentele rechten en fundamentele vrijheden. Het is dus niet verrassend dat deelname in biobanken niet vanzelfsprekend is. Daar waar sommige mensen graag hun gegevens bijdragen aan de wetenschap of gezondheidszorg, en zo inzicht geven in hun persoonlijke leven of levenssfeer, doen anderen dit bij voorbaat niet. Bijvoorbeeld, omdat zij zich zorgen maken om misbruik of veiligheid van hun gegevens bij een deelname. Bovendien veranderen sommigen die deelnemen van gedachte en trekken zich terug van deelname. Een belangrijke vraag is daarom: hoe kunnen we beter begrijpen wat de verschillen zijn tussen deelname/niet-deelname in onderzoeksgegevensverzamelingen en daarbij hun onvermijdelijke gevolgen.

Voorgaande onderzoeken laten zien dat verschillende vormen van maatschappelijk vertrouwen, zoals vertrouwen in andere individuen of vertrouwen in gezondheidszorg instituten, cruciaal zijn voor het begrijpen van deelname/niet-deelname. Daarom onderzoeken we, in dit proefschrift, de rol van vertrouwen bij deelname/niet-deelname waarbij we vooral focussen op vertrouwen in de maatschappij (i.e. vertrouwen in de overheid instituten en andere individuen), vertrouwen in onderzoek en vertrouwen in het delen van data. Het doel van dit proefschrift is om zo bij te dragen aan een beter begrip van deelname/niet-deelname in gegevensverzamelingen van onderzoek. Dit doen we door de context van medische wetenschappen, vooral epidemiologie, te gebruiken en ons te richten op deelname/niet-deelname in biobanken en grote cohort studies. Aangezien juist dit goede voorbeelden zijn van grootschalige, gecentraliseerde gegevensverzamelingen met gevoelige gegevens. Met dit als basis kunnen we vervolgens de gedrags- en morele mechanismen die een rol spelen ontdekken en begrijpen vanuit de ethiek, psychologie en sociologie.

De hoofdvraag van dit proefschrift is: *Welke factoren beïnvloeden de bereidheid van individuen om hun persoonlijke gegevens te delen in grootschalige, gecentraliseerde gegevensverzamelingen voor wetenschappelijk onderzoek, in het bijzonder biobanken?*

We beantwoorden deze vraag door eerst twee verkennende kwalitatieve studies te doen naar hetgeen dat samenhangt met deelname/niet-deelname in een biobank. We hielden hiervoor gesprekken met deelnemers en ex-deelnemers van een algemene Nederlandse bevolkingsbiobank, en met niet-deelnemers die hun gegevens niet bijdroegen aan deze biobank. Daarna voerden we drie kwantitatieve vervolgstudies uit om onze inzichten uit die gesprekken te

controleren en de relevante mechanismen en determinanten van deelname/niet-deelname nog beter te begrijpen.

Het eerste deel van ons onderzoek beschrijven we in Hoofdstuk 2 en 3. Deze hoofdstukken beschrijven de resultaten van de interviews met individuen. We verkennen en bespreken waarom sommigen wel of niet bijdragen aan een Nederlandse bevolkingsbiobank als voorbeeld van een grootschalige, gecentraliseerde gegevensverzameling. We onderzoeken en vergelijken hoe deelnemers en niet-deelnemers dachten over bijdragen aan zulk onderzoek en welke reden ex-deelnemers hadden om te stoppen met deelname. Bovendien geven we aan waarom en onder welke voorwaarden individuen een Nederlandse bevolkingsbiobank vertrouwen met hun persoonlijke gegevens. Zodoende laten we ook zien welke elementen afbreuk doen aan dit vertrouwen.

Het tweede deel van ons onderzoek is beschreven in Hoofdstukken 4 tot en met 6. Deze hoofdstukken gaan over vragenlijststudies. We vergeleken eerst in een vragenlijststudie antwoorden en gegevens van deelnemers en individuen die deelname aan een Nederlandse bevolkingsbiobank weigerden of zouden weigeren. De respondenten waren zodoende deelnemer van de Lifelines biobank of weigeraar van deelname aan Lifelines. Daardoor kregen we inzicht in deze twee groepen respondenten hun verschillen, met name door te kijken naar hun demografische en pro-sociale intra-persoonlijke kenmerken. Vervolgens analyseerden we de motieven van deelnemers hun deelname of een mogelijke stop van deelname. Daarbij keken we hoe de motieven samenhangen met verschillende kenmerken van de deelnemer. Tot slot keken we hoe individuen hun bijdrage van gegevens aan een Nederlandse bevolkingsbiobank samenhang met vertrouwen in de maatschappij en andere maatschappelijke bijdragen (bijvoorbeeld vrijwilligers werk of orgaandonatie). In deze laatste studie keken we daarom vooral naar de relatie tussen sociaal kapitaal en deelname in publieke goederen, terwijl we gebruik maakten van gegevens van het Centraal Bureau van Statistiek die gekoppeld werden met gegevens van de Lifelines biobank.

In het algemeen, geven de resultaten van dit proefschrift aan dat individuen hun mate van vertrouwen en pro-sociale waarden, tot op zekere hoogte, deelname aan een bevolkingsbiobank kunnen verklaren. Hoofdstuk 2 stelt dat de beslissing van zowel deelnemers als niet-deelnemers over hun bijdrage aan een biobank gemotiveerd is om anderen te helpen. Daar waar deelnemers voordelen zien en weinig risico's, maken niet-deelnemers en ex-deelnemers zich zorgen om persoonlijke of maatschappelijke schade. Onze analyse laat zien dat maatschappelijk vertrouwen, zoals vertrouwen in de overheid en andere individuen, en pro-sociale waarden – hier gedefinieerd als het betrekken van anderen (hun belang) – de verschillende visies van deelnemers en niet-deelnemers gedeelteli-

jk uitleggen.

Hoofdstuk 3 voegt aan deze resultaten toe dat de verschillen in mate van vertrouwen afhangt van vier soorten individuele overtuigingen. Ten eerste, de overtuiging dat grootschalige, gecentraliseerde gegevensverzameling voordeel geeft aan het algemeen publiek of de publieke zaak. Ten tweede, de overtuiging van een goede organisatie van de gegevensverzameling en opslag, bijvoorbeeld veilige verzameling van gegevens, de passende schaal van de gegevensverzameling, de goede reputatie of betrouwbaarheid van de verantwoordelijke organisatie en de motieven, competentie en het gedrag van diens medewerkers. Ten derde, de overtuiging dat de gegevens of inzichten van de verzamelde gegevens gevoelig zijn, afhankelijk van reikwijdte en context van verzameling. Ten vierde, de overtuiging van het toepassen van geïnformeerde toestemming als middel om controle te houden over persoonlijke gegevens. Al deze overtuigingen bleken invloed te hebben op vertrouwen en uiteindelijk op de beslissing om wel of niet te deelnemen, of te stoppen met deelname.

Onze kwantitatieve studies bevestigden onze resultaten uit de kwalitatieve studies en bouwden daar op voort. Hoofdstuk 4 laat zien dat gezondheids-gerelateerde waarden een belangrijke rol spelen in beslissingen over deelname, naast de mate van maatschappelijk vertrouwen. De drempel om deel te nemen aan Lifelines was bepaald door lage gezondheids-gerelateerde waarden en mate van maatschappelijk vertrouwen. Ondanks dat studies laten zien dat redenen voor onderzoek deelname pro-sociaal zijn, vonden wij niet dat deelnemers, vergeleken met weigeraars, anderen hun belang in het algemeen meer betrekken, als wij controleerden voor andere waarden en mate van vertrouwen. Noch vonden wij een hoger vertrouwen in onderzoekers buiten de context van biobanking. Hoofdstuk 5 bevestigt dat motieven voor deelname en stopzetting van deelname voortvloeien uit voordelen en schade voor de maatschappij en individuen zelf. Voorspellen van deelnemers hun bereidheid om gegevensuitwisseling voor medisch onderzoek toe te staan, zijn naast hoge mate van vertrouwen, een sterke motivatie om individuele en vooral maatschappelijke voordelen te behalen met deelname. Ook voorspelt daarbij een lage mate van bezorgdheid over individuele en maatschappelijke schade door deelname. Hoofdstuk 6 vertelt ons dat wanneer een individu een hoge mate van betrokkenheid in de maatschappij en vertrouwen in de maatschappij, ook bekend als sociaal kapitaal, toont, de kans hoger is dat die individu meedoet aan zowel Lifelines als een andere gegevensverzamelingen voor onderzoek. Die betrokkenheid in de maatschappij zien we vooral terug in het formele pro-sociale gedrag, zoals stemgedrag of lidmaatschappen van vrijwilligers organisaties.

Al met al laten de resultaten van dit proefschrift zien dat er een dilemma is

tussen het bijdragen aan de maatschappij zonder voorbij te gaan aan individuele interesses en zorgen. Terwijl deelnemers hoge pro-sociale waarden en een algemeen vertrouwen in de maatschappij (i.e. overheid en onderzoekers uit het publieke domein) moeten hebben voor deelname aan gegevensverzamelingen voor medisch onderzoek, vraagt het besluit om deelname ook om de waardering van en de ervaring in de specifieke context van gegevens verzameling en onderzoek, bijvoorbeeld gezondheid en gezondheidszorg. Ex-deelnemers gaven inzicht in hoe hun positieve houding tot deelname veranderde door een groeiend tekort aan maatschappelijk vertrouwen of specifiek vertrouwen in de biomedische context. Het onderzoek van dit proefschrift is daarmee één van de eerste dat aan geeft hoe vertrouwen de bereidheid tot deelname en vooral die van het stopzetten van deelname, kan beïnvloeden afhankelijk van overtuigingen met betrekking tot de maatschappij en de waardering van de context van de gegevensverzameling.

Dit proefschrift onderstreept de diversiteit van motieven van deelname in gegevensverzamelingen en het wederkerige karakter ervan. Onze bevindingen laten zien dat de beslissing van deelnemers en niet-deelnemers om wel of niet bij te dragen aan een biobank, tot op zekere hoogte, is vanwege de wil om anderen te helpen, in het bijzonder toekomstige generaties. We vonden dat de beslissing tot deelname in gegevensverzamelingen voor gezondheidsonderzoek samenhangt met pro-sociale houdingen en pro-sociale gedrag buiten de biomedische context, zoals vrijwilligerswerk of stemgedrag. Dit kan er op wijzen dat het publieke belang belangrijk is voor deelname. Echter gaat deelname niet alleen om het helpen van anderen. Onze resultaten laten zien dat lage maten van algemene pro-sociale houdingen en oriëntaties of andere pro-sociale motieven, niet gelijk is aan niet-deelname. Gezondheid gerelateerde waarden hingen onafhankelijk samen met deelname in gegevensverzamelingen, wat het belang van individuele interesses toont. Daarnaast waren er nog twee andere factoren omtrent de individuen die een rol spelen in het beslissen over deelname, namelijk de ervaring in onderzoekdeelname van een individu en de ervaringen van hun sociale netwerk.

Het zou kunnen om individuen te overtuigen om deel te nemen aan onderzoek, door andere motieven aan te spreken dan het helpen van anderen in de maatschappij. Bijvoorbeeld, door het aanspreken van motieven om anderen binnen hun sociale netwerk te volgen, of om individuele interesses te vervullen. Door het benadrukken van de kracht van gegevensverzamelingen zouden meer verschillende individuen geworven kunnen worden voor deelname. Die kracht ligt zowel in de toegevoegde waarde van aanvullende informatie uit gegevensverzamelingen als in mogelijke individuele voordelen van deelnemers. Dit kan

misschien leiden tot een groeiend aantal mensen met ervaring in deelname in gegevensverzamelingen voor onderzoek. Ook zouden er meer deelnemers uit verschillende sociale netwerken kunnen komen. Uiteindelijk zou dit alles zorgen voor meer deelname in gegevensverzamelingen. Vooral van groepen die juist op dit moment vaak niet voor komen in onderzoek en gegevensverzamelingen.

Ondanks dat vertrouwen in de maatschappij en vertrouwen in de relevante context van gegevensverzameling sterk gecorreleerd zijn, verklaarden zij unieke variantie in deelname/niet-deelname. Deelnemers hadden meer vertrouwen in de overheid en in andere individuen. Tegelijkertijd waren mensen die in het algemeen meer vertrouwen hebben in andere individuen vaker deelnemer in meerdere gegevensverzamelingen. Desondanks hadden deelnemers met een hogere mate van maatschappelijk vertrouwen of vertrouwen in anderen niet meer vertrouwen in commerciële bedrijven. Zo vonden deelnemers maatschappelijke schade niet een belangrijkere reden voor het stopzetten van hun deelname. Wel vonden zij individuele voordelen belangrijker en hadden zij minder zorgen om individuele schade. Dit laat voor het beslissen over deelname zien hoe belangrijk overtuigingen zijn van een positief gevolg van onderzoek, vertrouwen in onderzoek en de betrouwbaarheid van de context van onderzoek.

Om de kans te vergroten dat individuen mee willen doen aan grootschalige, gecentraliseerde gegevensverzamelingen die gegevensuitwisseling toepassen, is vertrouwen in de publieke en private sfeer nodig als brandstof voor het besluit tot deelname. Zowel de mogelijkheid als de uitdaging in het bewaken van deelname in gegevensverzamelingen voor toekomstig innovatief onderzoek is noodzaak voor hoge mate van vertrouwen in het algemeen. Het zou daarom cruciaal kunnen zijn om individuen te informeren over de mogelijkheden en onmogelijkheden van gegevensverzamelingen vanuit een breed perspectief. Bijvoorbeeld, door informatie te geven over de hele context van onderzoek en mogelijke voordelen en schade op individueel en maatschappelijk niveau. Dit kan het gevoel van controle over data vergroten van mogelijke deelnemers en verhelderen wanneer welke contextuele normen gelden, terwijl het vertrouwen versterkt wordt.

Tegenwoordig kunnen biobanken en andere gegevensverzamelingen grenzen en contexten overschrijden via complexe systemen, vooral wanneer zij gegevens uitwisselingen en sensor technologieën toepassen. Dit heeft tot gevolg, dat de verantwoordelijkheden van grootschalige, gecentraliseerde gegevensverzamelingen misschien meer liggen op een systeem niveau, zoals informatie en toezicht structuren, dan op individueel niveau, zoals het verkrijgen van individuele geïnformeerde toestemmingen of het maken van een veilige gegevensverzameling. De focus van gegevensverzamelingen in wetenschappelijk onderzoek

zou daarom moeten verschuiven naar systemen voor het verkrijgen en toezicht houden op data die proactief geïnformeerde besluitvorming over deelname mogelijk maken. Dit gaat dus verder dan alleen het bevestigen dat een deelnemer voldoende kennis heeft over nadelen en voordelen van het betreffende onderzoek via een geïnformeerd toestemmingsformulier. Noch beperkt het zich tot reageren op incidenten zoals data lekken.

Organisaties die verantwoordelijk zijn voor gegevensverzamelingen zouden er voor moeten zorgen dat toezichtsystemen en consequenties van deelname/niet-deelname in onderzoek duidelijk en continu gecommuniceerd worden. Dit moet enerzijds met deelnemers en anderzijds met anderen in de maatschappij, met name voor bevolkingsgerichte gegevensverzamelingen. Een geïnformeerde toestemming kan dan echt betekenen dat iemand geïnformeerd is waarvoor toestemming wordt gegeven en hoe deze toestemming aangepast kan worden. Bovendien wordt het dan helder wanneer en hoe je wordt geïnformeerd over situaties die om een vernieuwing of aanpassing van je gegeven toestemming vragen. Een stap in de goede richting zou kunnen zijn dat onderzoekers mogelijke deelnemers hun begrip van data impact vergroten door visualisaties van data processen te gebruiken.

Niet alleen organisaties die een gegevensverzameling opzetten zijn verantwoordelijk om te zorgen dat beslissingen tot deelname geïnformeerd zijn en gebaseerd op feiten. Overheden en publieke instellingen die onderzoek op nationaal en internationaal niveau vertegenwoordigen zijn cruciaal in een robuust toezicht op onderzoeksgegevensverzamelingen. Deze kunnen en zouden moeten helpen met het informeren, bediscussiëren en besluiten over verantwoordelijkheden en risico's op mogelijke maatschappelijke en individuele schade door gegevensverzamelingen. Nationale en internationale autoriteiten van toezicht, zoals data beschermingsinstituten en ethische toetsingscommissies zouden ook de mogelijkheden en beperkingen van gegevensverzamelingen verder moeten verhelderen. Daarnaast zouden ze het proactieve toezicht op gegevensverzamelingen en hun gebruik moeten versterken. Door een beter overzicht te creëren van beschikbare gegevens en gegevensverzamelingen kan het publiek meer duidelijkheid en ondersteuning krijgen. Wat weer bij kan dragen aan een efficiënter gebruik van gegevensverzamelingen. Het in leven roepen van data

principes zoals *FAIR (Findable, Accessible, Interoperable and Reusable)* of data infrastructuur initiatieven als *HEALTH-Research Infrastructure* is ook een stap in de goede richting om inzicht en overeenstemming te krijgen van mogelijkheden en beperkingen van gegevensverzamelingen.

Conclusie

Concluderend, waardes en vertrouwen bieden veel mogelijkheden om besluiten van deelname of niet-deelname, of het stopzetten van deelname in gegevensverzamelingen voor onderzoek beter te begrijpen. Als deelname beslissingen worden gemaakt met verwachtingen die in lijn zijn met de werkelijkheid en gewaarborgd door sterke toezicht structuren, dan doen deze beslissingen een breed vertrouwen in de maatschappij eer aan. Het is niet zo zeer het feit *dat* een vraag naar gegevens wordt gedaan, maar *hoe* deze vraag wordt gesteld en *wat* de voorwaarden zijn van dit verzoek om gegevens. We hopen dat onze bevindingen bijdragen aan praktische interventies, die gericht zijn op het juist benaderen van mogelijke deelnemers hun ethische en praktische zorgen. Om zo ook het vertrouwen en de kans op deelname in gegevensverzameling voor wetenschappelijk onderzoek te onderhouden en te vergroten. Op deze manier helpt dit onderzoek samen met vervolg onderzoek om verder rijke, grootschalige gegevensverzamelingen van gezondheid gerelateerde en andere maatschappelijke relevante onderwerpen.

Kokkuvõte

Usalduse olulisus ja teadusandmebaasidega liitumine

Viimase kolmekümne aasta jooksul on teaduslikel eesmärkidel kogutud isikuandmete säilitamine paberkandjal asendumas digitaalsete teadusandmebaasidega. Selline digitaliseerumine on lihtsustanud andmete genereerimist, hoiustamist, kasutamist ja vahetamist ning seeläbi muutnud digitaalsed teadusandmebaasid üheks peamiseks andmeallikaks teadusuuringutes. Tehnoloogia arengu tõttu suureneb pidevalt andmete produtseerimise võimekus ning avanevad võimalused uut tüüpi teadusuuringuteks, mis omakorda kasutavad suuremahulisi tsentraliseeritud digitaalseid teadusandmebaase.

Biomeditsiin on teadusvaldkond, mida sellised uudsed uuringuvõimalused mõjutavad, kuna suurandmed (big data) omavad suurt tähtsust tervise ning tervisekäitumise paremal mõistmisel ja edendamisel üksikisiku tasandil. See aga eeldab, et teadusandmebaasidega on nõus oma isikuandmeid jagama palju inimesi, mis muudabki andmebaasid suuremahulisteks. Sellistes teadusandmebaasides säilitatakse suurtes kogustes nii bioloogilist materjali kui korduvküsitlustel ja -mõõtmistel põhinevaid isikuandmeid pikema ajaperioodi kohta. Biomeditsiinis on üheks selliseks andmebaasiks biopank.

Olakse üksmeelel, et digitaalsetele teadusandmebaasidele üleminekuga kaasnevad suured muudatused, mis on eriti seotud tundlike andmete käsitlemisega. Digitaalsete teadusandmebaaside abil loodavad uued teadmised võivad tuua kasu nii üksikisikule kui kogu ühiskonnale, kuid võivad tekitada ka uusi ohte, mistõttu pole suurte teadusandmebaasidega liitumine sugugi iseenestmõistetav. Osad inimesed on nõus oma isikuandmeid jagama eesmärgiga panustada teadusesse ja saada seeläbi rohkem teadmisi ka enda kohta, kuid teised muretsevad oma isikuandmete turvalisuse pärast, seda juhul kui nad liituksid teadusandmebaasidega. Mõned inimesed aga on pärast järelemõtlemist valmis muutma oma otsust andmete jagamise/mittejagamise osas. Seega on oluline paremini mõista, mille poolest kõik need inimesed erinevad.

Varasemad uuringud on näidanud, et usaldus teiste inimeste ja tervishoiuasutuste suhtes võivad lahti seletada inimeste soovi/soovimatust oma andmeid teadusandmebaasidega jagada. Käesoleva väitekirja eesmärk oli aru saada teadusandmebaasidega liitumise/mitteliitumise otsustest keskendudes usaldusele ühiskonna (riiklikud institutsioonid, teised inimesed), teaduse ja isiklike andmete jagamise vastu. Töös kasutati kvalitatiivseid ja kvantitatiivseid epidemioloogilisi meetodeid, et uurida biopankade ja ulatuslike kohortuuringutega liitumise/mitteliitumise eetilisi, sotsiaalseid ja psühhosotsiaalseid arusaamu

ning sellega seotud käitumis- ja moraalimehhanisme.

Vastused väitekirja peamisele uurimisküsimusele „Missugused tegurid mõjutavad üksikisiku soovi teaduslikel eesmärkidel jagada oma andmeid suuremahuliste tsentraliseeritud teadusandmebaasidega nagu biopangad?“ põhinesid kahel kvalitatiivsel uuringul, mille käigus viidi läbi süvaintervjuud Hollandi rahvastikupõhise biopangaga liitunute, liituda soovijate ja mittesoovijatega ning kolmel kvantitatiivsel uuringul, mille käigus uuriti biopangaga liitumise soovimise/mittesoovimise mehhanisme ja mõjureid.

Töö tulemused näitasid, et otsus anda või mitte anda oma panus biopanka võis tuleneda soovist uute tõenduspõhiste teadmistega aidata teisi inimesi, aga eelkõige toetada tulevase põlvkondi. Kuid teadusandmebaasiga liitumise soov ei olnud seotud vaid teiste aitamise eesmärgiga, sest prosotsiaalsete motiivide puudumine ei tähendanud liitumise soovi puudumist. Tervisega seotud väärtused olid sõltumatult seotud teadusandmebaasidega liitumisega, mis näitas individuaalse huvi olulisust. Lisaks üldisele huvile head teha ning individuaalsele huvile tervise ja tervishoiu vastu, mõjutasid liitumise otsust sõltumatult teistest teguritest veel kaks individuaalse huviga seotud tegurit nagu üksikisikute teadustöös osalemise kogemus ja oma sotsiaälvõrgustiku mõju.

Üksikisikuid võib olla võimalik veenda käituma prosotsiaalselt ja liituma teadusandmebaasidega ilma ühiskonnas teiste inimeste abistamise motiivideta. Näiteks kui motiiviks on järgida oma sotsiaälvõrgustikus kehtivaid norme või järgida individuaalset huvi. Teadusandmebaasidesse kaasatus võiks suurenedagi veelgi kui rõhutada nii andmebaasile lisaväärtuse tekitamist iga üksikisiku liitumisel kui sellest saadavat võimalikku kasu üksikisikule (nt korrapärane tervisekontroll).

Käesoleva töö tulemusena leiti, et usaldus ühiskonna vastu oli tugevas korrelatsioonis huviga andmebaasidel põhinevate teadusuuringute konteksti vastu, samas ennustasid mõlemad sõltumatult liitumise/mitteliitumise soovi. Potentsiaalsed teadusandmebaasidega liitujad ei muretsenud võimaliku sotsiaalse kahju pärast, küll aga tõid nad esile suurema individuaalse kasu motiivi. Need tulemused näitavad, kui oluline roll on positiivsel suhtumisel teadusuuringutesse ja teadusuuringu konteksti usaldusväärsusel teadusandmebaasidega liitumise otsuse langetamisel. Seega peaks üksikisikute suuremahulistesse tsentraliseeritud teadusandmebaasidesse kaasamise suurendamiseks olema piisav usaldus avaliku ja erasektori vastu, mida peaks mõistma kui võimalust ja väljakutset tulevaste uuenduslike teadusuuringute läbiviimiseks ja teadusandmebaasidega liitunute kaitsmiseks. Seetõttu on teadusandmebaasidega liitumisel võtmeküsimuseks üksikisikute teavitamine andmebaasidega liitumise tugevustest ja nõrkustest ning iga teadusuuringu konteksti kohta põhja-

liku teabe jagamine koos võimaliku kasu ja kahju kirjeldamisega nii üksikisiku kui ühiskonna tasandil. See suurendab potentsiaalsete liitujate tajutud kontrolli isikuandmete üle, suurendades samal ajal usaldust ning selgitades, millised kontekstuaalsed normid kehtivad ja millal.

Tänapäeva keerukad tehnoloogilised lahendused biopankades ja teadusandmebaasides nagu näiteks andmete linkimine ja sensorandmete tehnoloogiad, võimaldavad lihtsasti läbi viia interdistsiplinaarseid ja rahvusvahelisi uuringuid. Seetõttu peab teadusandmebaaside eest vastutamine toimuma süsteemselt, jagades teadusandmebaasiga liitujatele ja teistele ühiskonna liikmetele pidevalt selget infot teadusuuringutes osalemise/mitteosalemise tagajärgedest, seda just eriti rahvastikupõhiste teadusandmebaaside korral. Oluliseks sammuks on tegelike või potentsiaalsete liitujate teadlikkuse suurendamine andmete jagamisega seotust, rakendades selleks rohkem andmeprotsesside visualiseerimist teadlaste endi poolt.

Lisaks andmebaaside haldajatele on teadusuuringuid riiklikul ja rahvusvahelisel tasandil esindaval valitsusel ja asutustel teadusuuringute andmebaaside süsteemses juhtimises võtmetähtsus. Nad peaksid aitama andmebaasidega kaasneva võimaliku ühiskondliku ning individuaalse kasu ja kahjuga seotud kohustustest ja riskidest teavitamisel, nendest arutamisel ja nende üle otsustamisel. Riiklikud ja rahvusvahelised institutsioonid nagu andmekaitsega seotud asutused ja eetikakomiteed peaksid esile tooma teadusandmebaaside tugevaid ja nõrku külgi ning tugevdama teadusandmebaaside andmete edasise kasutamise proaktiivset järelevalvet. Parema ülevaate tagamine olemasolevatest andmetest ja teadusandmebaasidest annab üldsusele rohkem juhiseid ja selgust ning soodustab teadusandmebaaside tõhusat kasutamist. Selliste andmepõhimõtete nagu *FAIR* (*Findable, Accessible, Interoperable, Reusable*) või teadusuuringute infrastruktuuride algatuste nagu *HEALTH-Research Infrastructure* kehtestamine oleks oluline samm edasisel teadusandmebaasidega seotud võimaluste ja piirangute paremal mõistmisel ning nende suhtes üksmeele saavutamisel.

Järeldused

Kokkuvõtteks võib öelda, et väärtused ja usaldus annavad olulise aluse teadusandmebaasidega liitumise/mitteliitumise või liitumisest loobumise otsuste mõistmiseks. Kui liitumisotsused tehakse ootustega, mis on vastavuses tege­likkusega ja toimuvad põhjalike juhtimisstruktuuride raames, siis need näitavad laiapõhjalist usaldust ühiskonna vastu. Oluline on mitte see, et esitatakse uuringus osalemise kutse, vaid pigem see, kuidas indiviidi poole pöördutakse ja millised on osalemise tingimused.

Väitekirja tulemused on aluseks praktilistele sekkumistele, mis on suunatud võimalike teadusandmebaasidega liitujate eetiliste ja praktiliste murede nõuetekohaseks lahendamiseks, nende usalduse säilitamiseks ja tõstmiseks ning selleks, et suurendada oma isikuandmete teadusuuringutele jagamise tõenäosust. Sel viisil aitab antud väitekirja koos selle valdkonna järeluurimistega kaasa suuremahuliste teadusandmebaaside edasisele laialdasele levikule ja kasutusele, et leidavastuseid uutele tervishoiu ja muude eluvaldkondadega seotud küsimustele.

Research Institute SHARE theses

This thesis is published within the Research Institute SHARE (Science in Healthy Ageing and healthcaRE) of the University of Groningen / the University Medical Center Groningen. Further information regarding the institute and its research can be obtained from our internet website: <https://umcgresearch.org/nl/web/research/w/share>.

Information about more recent theses can be found in the list below.

2021

Peters RM

Total hip arthroplasty: register based evaluation of current practice
(prof SK Bulstra, dr M Stevens, dr WP Zijlstra)

Simmelink L

Combined arm-leg ergometry in persons with a lower limb amputation
(prof R Dekker, prof JHB Geertzen, prof LHV van der Woude)

Valk TA

When multiple become one; understanding the coordination of degrees of freedom in synergies in goal-directed pointing movements
(dr RM Bongers, prof E Otten, dr LJ Mouton)

Bos FM

Ecological momentary assessment as a clinical tool in psychiatry: promises, pitfalls, and possibilities
(prof M Wichers, prof R Bruggeman, dr JAJ van der Krieke, dr E Snippe)

Akbari F

Chronic pain in the context of the lives of dyads; cognitions, behaviors, and well-being

(prof M Hagedoorn, prof R Sanderman, dr M Dehghani)

Hepping AM

Grip on recovery after paediatric forearm fractures

(prof SK Bulstra, prof JHB Geertzen, dr M Stevens)

Beijers L

Parsing the heterogeneity of major depression

(prof RA Schoevers, dr KJ Wardenaar, dr HM van Loo)

Köhler TC

Providing color to the pharmacy technician; a new profession within the pharmacy team

(prof ADC Jaarsma, dr M Westerman)

Bunk SF

Frontal brain functioning and pain; possible underlying mechanisms of increased pain responses in age- and dementia-related cognitive impairment

(prof SU Zuidema, dr M Kunz)

Bosáková L

Breaking the cycle of poverty; routes to counteract intergenerational transmission of socioeconomic health differences

(prof SA Reijneveld, prof A Madarasová-Gecková)

Vendeloo SN van

Optimizing learning environments and resident well-being in postgraduate medical education

(prof PLP Brand, prof SK Bulstra, dr CCPM Verheyen)

Sampurna MTA

Improving the management of hyperbilirubinemia in a limited-resource area

(prof AF Bos, dr CV Hulzebos, dr PH Dijk, dr R Etika)

Siswanto JE

Retinopathy of prematurity; how to prevent retinopathy of prematurity in

preterm infants in Indonesia?
(prof AF Bos, prof A Adisasmita, dr PH Dijk)

Kaper MS

Improving communication in healthcare for patients with low health literacy; building competencies of health professionals and shifting towards health literacy friendly organizations
(prof SA Reijneveld, prof AF de Winter)

Schuurmans J

Population-based expanded carrier screening reporting couple results only: a mixed methods approach
(prof IM van Langen, prof AM Lucassen, prof AV Ranchor, dr M Plantinga)

Vrijsen J

Genetics and epidemiology of glaucoma and myopia
(dr N Smidt, prof SEJA de Rooij)

Asefa, NG

Towards dementia risk reduction among individuals with a parental family history of dementia
(prof H Snieder, prof NM Jansosius)

Tegegne, BT

The Heart of the Matter: Discovery of new genetic loci for heart rate variability and its relationship with blood pressure and mortality
(prof H Snieder, dr H Riese, dr IM Nolte)

For information about earlier theses, please, visit our website.

Acknowledgments

Since there are too many of you, who have provided me the invaluable scientific and moral support during my PhD journey, I find it difficult to name you all. Hence, I will do my best to express my gratitude and view of this joyful journey. My PhD research was like sailing on a ship to an unknown, yet presumably better and even more exiting, destination. This destination would have been impossible to reach without a supportive “crew” on my side, since every ship requires several hands to sail. I would describe this ship as a Frisian ‘Skûtsje’, which requires more hands and skills than average. I could only arrive to my destination with the help of a large crew, everyone having their unique and unmeasurable input to this journey. Some of you accompanied me for a long expedition, while others joined in for short moments, but had the best possible timing to keep this ship on the right course. Nevertheless, all of you played a crucial role in their own manner, and during a particular moment in time. I am grateful for all who spend time with me during this journey, and I hope this collaboration helped them to reach a better destination as well. Yes, I do TRUST that becoming Dr. R. Broekstra provides me with better chances in life.

The honour to work and learn from my four (co-)promotors, all with different specialization, has helped me to broaden my horizon. I have enjoyed working with several project collaborators, having supportive colleagues from three different departments (Department of Epidemiology, of Social Psychology, and of Ethics), and relying on old and new friends from different countries. I find it hard to make a trustworthy distinction between my private and working life since these are heavily intertwined. For me, the science became a hobby, which frequently got out of my hands, which has its advantages and disadvantages. Yet, I will try to express my gratitude, and trust you to understand my specific level of gratitude for you.

Colleagues, friends, and the valued nameless contributors. Thank you for your valuable input and support for those uncountable times. You helped me by getting me on track, or keeping me on track to obtain this PhD. With your help I found the bits and pieces of the researcher, scientist, and the non-scientist in me, who developed through the time thanks to you. This is why I always felt at ease despite exploring unfamiliar waters, and understand what the feeling of ‘at home’ is about while surrounded by the right people.

My promotores, family, and close friends. Thank you for your guidance, support, understanding, and thought provoking discussions during my PhD journey. As a real “crew” you always stood by my side, even though I sometimes forgot or did not yet know how to express my gratitude. We shared moments of joy which equipped us with good memories and stories to remind when we look back on this journey taken together. In difficult times, you helped me to navigate through unknown waters without letting the ship to sink. Your support made me feel safe, so the unknown waters became known. Without your support, this PhD would not have been so memorable and fun, therefore: dankjewel, danke, dankjewol, thank you, hvala vam, aitäh. I hope you will value these words, and that you will value our shared memory as much as I do! The only wish for my future is to share even more moments together, so let me know when we can hoist the sails for another journey.

Katri. Even though I am rarely in lack of words, I am for this part. Though obtaining a PhD degree is known to improve the chances in life, I could never have expected its implications for my life. This ‘wedding’ with academia is meaningless without you by my side. Katri, suur aitäh Sulle selle eest, et Sa kuulud minu ellu. Sinuga koos tunnen ma ennast õnnelikuna igas maailma nurgas.

About the author

Biography

Reinder Broekstra was born on February 8, 1988 in Leeuwarden, The Netherlands, and raised in the port city of Harlingen, The Netherlands. His multidisciplinary interest started during his grammar school and eventually resulted in graduating from a Bachelor of Psychology and a Bachelor Laws at the University of Groningen. He started his research career with conducting applied research in the commercial sector. After finishing his master Psychology at that same university in 2016, he continued with research conducting research in an interdisciplinary scientific project called “Big data and the dilemma of innovative knowledge versus threats to personal integrity” funded by the Dutch Research Council with file number 313-99-313. The successful collaboration in this project enabled him to obtain a PhD degree. Currently, he conducts research into ethical legal and (psycho)social issues in genetics and genomics at the University Medical Center Groningen and the University of Southampton. His key interests and expertise are in topics related to the implications of innovation in information technology and participation/non-participation in or willingness/non-willingness to contribute to research, especially in the medical scientific field.

Curriculum Vitae

List of scientific publications

Vrijzen, J., Maeckelberghe, E. L. M., Broekstra, R., de Vries, J. J., Abu-Hanna, A., De Deyn, P. P., Voshaar, R. C. O., Reesink, F. E., Buskens, E., de Rooij, S. E., & Smidt, N. (2021). Knowledge, health beliefs and attitudes towards dementia and dementia risk reduction among descendants of people with dementia: a qualitative study using focus group discussions. *BMC Public Health* 21:1, 21(1), 1–10. <https://doi.org/10.1186/S12889-021-11415-2>

Broekstra, R., Aris-Meijer, J., Maeckelberghe, E., Stolk, R., & Otten, S. (2021). Demographic and prosocial intrapersonal characteristics of biobank participants and refusers: the findings of a survey in the Netherlands. *European Journal of Human Genetics*, 29(1), 11–19. <https://doi.org/10.1038/s41431-020-0701-1>

Broekstra, R., Maeckelberghe, E. L. M., Aris-Meijer, J. L., Stolk, R. P., & Otten, S. (2020). Motives of contributing personal data for health research: (non-)participation in a Dutch biobank. *BMC Medical Ethics*, 21(1), 62. <https://doi.org/10.1186/s12910-020-00504-3>

Broekstra, R., Aris-Meijer, J., Maeckelberge, E., Otten, S., Stolk, R., Maeckelberghe, E., Stolk, R., & Otten, S. (2020). Trust in Centralized Large-Scale Data Repository: A Qualitative Analysis. *Journal of Empirical Research on Human Research Ethics*, 15(4), 365–378. <https://doi.org/10.1177/1556264619888365>

Broekstra, R., Maeckelberghe, E. L. M., & Stolk, R. P. (2017). Written informed consent in health research is outdated. *European Journal of Public Health*, 27(2), 194–195. <https://doi.org/10.1093/eurpub/ckw198>

Work experience - Research

sep '20 - present	Postdoctoral researcher, University Medical Center / University of Groningen, University of Southampton Responsible for research projects relating to (potential) Ethical, Legal, Social of implications of innovative information technologies in medicine and medical sciences, especially clinical genetics and genomics.
feb '16 - dec '21	PhD-candidate, University Medical Center / University of Groningen Responsible for a research project about understanding perceptions of 'big data' and participation/non-participation in 'big data' research, which was funded by the Dutch Research Council program responsible innovation.

feb '15 - aug '15	Research-assistant PO&O, FPC Dr. S. van Mesdag Responsible for a research project, in combination with obtaining an master Psychology, that emphasized the organizational absenteeism policy and its effect.
jul '12 - jul '15	Owner and Independent Researcher Breinder Conducting qualitative and quantitative research for large national and international enterprises.
sep '14 - sep'15	Research-assistant department Criminal Law and Criminology, University of Groningen
jun '10 - jul '10	Intern Research-assistant Signicom: Market research (Amsterdam)

Education

feb '16 - dec '21	PhD: Medical sciences, University of Groningen
feb '15 - jan '16	Master Psychology: Industrial & Organisational, University of Groningen
sep '10 - jun '15	Bachelor of Law, University of Groningen
sep '07 - jun '14	Bachelor Psychology, University of Groningen
sep '00 - aug '07	Gymnasium, O.S.G. Piter Jelles Leeuwarden

Extracurricular activities

may '18 - dec '20	Vice-President PhD Graduates Network Netherlands
may '11 - may '16	Chairman Supervisory Board, Study Association Simon van der Aa
sep '11 - aug '13	General member Health and Safety Committee, Faculties of Law and Arts
sep '11 - aug '13	Group chairman, Progressief Rechten representative advisory body of the Faculty of Law
may '11 - may '12	Chairman Executive Board, Study Association Simon van der Aa