Opening the Personal Gate between Technology and Health Care
M. Giacomini et al. (Eds.)

doi:10.3233/SHTI260448

Trust, Beliefs, and Bias: Understanding Preferences for Consent Procedures for Secondary Use of Health Data

Melissa van Essen [a, b, 1], Robert Verheij [a, b, c], Isabelle Bos [b]
[a] *Tilburg School of Social and Behavioral Sciences, Tranzo, Tilburg, The Netherlands*
[b] *Netherlands Institute for Health Services Research, Nivel, Utrecht, The Netherlands*
[c] *Health Care Institute Netherlands, Diemen, Netherlands*
ORCiD ID: Melissa van Essen https://orcid.org/0009-0008-2299-1091, Isabelle Bos https://orcid.org/0000-0001-9648-6428, Robert Verheij https://orcid.org/0000-0003-2234-4819

Abstract. Secondary use of routine health data requires regulatory frameworks and consent procedures as a legal basis for data protection. The different consent models, however, influence representativeness, and thus quality, of the data. This study examines how socioeconomic factors, beliefs, and attitudes shape citizens' preferences for consent procedures, providing insight into potential biases in healthcare databases and the characteristics of patients who choose not to share their data. Using questionnaire data from 3,709 Dutch respondents, descriptive and forward selection regression analyses were conducted. Results show that predictors related to societal trust significantly influence preference for an opt-in procedure, while respondents with no preference for a consent procedure show similar patterns to those who are uncertain regarding their preference, involving trust, media, and politics. These findings are relevant to the proposed opt-out system within the European Health Data Space (EHDS), which may not align with citizens' preferences and may require active communication to gain public trust.

Keywords. Data quality, routine health data, consent procedures, trust, bias

1. Introduction

It is evident that routine healthcare data, like electronic health records (EHR), are increasingly being used for secondary purposes, such as scientific research, the development of healthcare products and policymaking purposes. [1-3] With this increasing use comes a growing need for clarity on how these data may be handled, leading to the adoption of legal and ethical protection measures including the General Data Protection Regulation (GDPR) and the European Health Data Space (EHDS). These regulatory frameworks ensure that the rights of data subjects are respected while promoting responsible access to healthcare data. [4, 5] Advancements like these, and their effects may significantly impact the quality of routine healthcare data.

When routine health data are concerned, the most common consent procedures in practice include primarily opt-in, and occasionally opt-out. [6] Each procedure has advantages and disadvantages and influences the consent rate, potentially leading to the introduction of consent bias and selection bias. [6, 7] Consequently, the choice of consent

---

[1] Corresponding author: Melissa van Essen, m.h.j.vanessen@tilburguniversity.edu

procedure strongly correlates with the representative aspect of and thereby with the reliability of study outcomes. [8] Understanding and minimizing these biases is a prerequisite for sound scientific research, underscoring the need to expand current knowledge of this issue and to translate findings into effective governmental action. Previous research shows beliefs and attitudes play a crucial role in the preferences of patients and their decision making, for example the preferences for specific consent procedures. [9, 10] These factors are highly dependent on cultural and political contexts and are therefore important to study in the Dutch context.

This study aims to answer the research question: are individual characteristics - including demographic and socioeconomic characteristics, as well as cultural and political beliefs and attitudes - associated with preferences for consent procedures and the willingness to share health data at all? The findings will contribute to the insight into the amount of bias in healthcare databases, and help researchers identify potential inclusion biases in their studies. This research will provide a deeper under-standing of patients who choose not to share health data and will help policymakers to address worries and distrust among citizens regarding the secondary use of healthcare data.

2. Methods

Data were collected using the LISS panel (Longitudinal Internet studies for the Social Sciences), managed by the non-profit research institute Centerdata (Tilburg University, the Netherlands). A questionnaire was sent to a representative, random sample of 4,818 Dutch panel members aged 16 years or older in December 2023. The questionnaire included multiple choice questions about the willingness to share healthcare data for secondary purposes, and the level of trust panel members have in several institutions that regularly use healthcare data for secondary purposes. 3,709 panel members completed the questionnaire in full (response rate 77.0%). Additionally, this study made use of the longitudinal LISS Core Study questionnaires of 2023 or 2024, and socio-economic, demographic, and other general background information on households and respondents were available through the LISS panel as well. [11] All 895 variables were linked together using the unique respondent ID variable. Variables without a conceptual link to the study aims were removed. Further selection of the final 92 variables was based on relevance to the study objectives and prior use in the literature. Variables were divided into different themes: Religion and ethnicity (9 items), activities (16 items), health (14 items), household (7 items), media (9 items), politics (12 items), trust (25 items). All variables were categorical or operationalized into a categorical variable. Variables with >50% missing values were excluded regardless of relevance, and variables with >5%<50% missing values were imputed. Of variables with <5% missing values the corresponding cases were excluded.

First, a descriptive analysis was performed on the demographic characteristics of the study population amongst the groups. Subsequently, we investigated which patient characteristics were associated with preferences for consent procedure types. Finally, we examined the characteristics of patients who reported to prefer not to share health data. To analyze which characteristics, beliefs and attitudes were associated with the outcome measures, two separate forward selection regression analyses were conducted.

Prior to the analyses, all reference categories were releveled and univariate screening was performed to further filter the relevant variables and achieve a well-defined selection of predictors. For the forward selection regression analyses, predictor variables were entered sequentially, and at each step the variable with the strongest association was added to the model. An entry criterion of $p < 0.05$ was handled, and variables were remained unless they failed to retain significance at $p < 0.10$, continuing until no further predictors met this inclusion threshold. The analyses were performed using R and RStudio, version 2023.06.1+561.

3. Results

Preliminary results show that 60% of the respondents prefer an opt-in procedure and 30% prefer an opt-out procedure. Approximately 6% of respondents have no preference and 4% don't know. There are no considerable differences in the age, sex or marital status of respondents between consent-groups. There are notable differences in socioeconomic status for the group that responded 'don't know' as they report no income or lower income more often, and reported 'don't know' regarding their income significantly more often. Respondents who were not willing to share their data (14%) reported 'don't know' regarding their income significantly more often as well. Respondents that responded 'don't know' regarding their consent system preference and respondents who were not willing to share their data had a primary school or pre-vocational secondary education more often, and a higher professional education or university education less often.

Preliminary results of the forward selection regression analyses show that, generally, variables with the theme 'trust' are statistically significant for respondents who prefer an opt-in consent procedure, compared to respondents who prefer opt-out. Table 1 shows that of the top 10 strongest significant predictors for the preference of an opt-in consent procedure, all 10 predictors belong to the theme 'trust'. The predictors for respondents who reported 'don't know' are similar to those of respondents who do not have a preferential consent procedure. Table 2 shows the overlapping predictors for those respondents, and predictors with the theme 'media' and 'politics' are statistically significant more often than for people who prefer opt-out or opt-in as a consent procedure.

Among the statistically significant variables, few belong to the theme 'health', or relate to the household and daily activities.

Table 1. Top 10 significant predictors for preference of opt-in consent procedure (reference category opt-out)

Theme	Predictor (for preference of opt-in procedure)	Category	OR	p-value
Trust	Provision of information regarding reuse of data	No permission to reuse	20.46	<0.001
Trust	Provision of information regarding reuse of data	Actively via a message, with every occurrence of reuse	5.14	<0.001
Trust	Trust in competent data reuse: scientific researchers	No trust	3.20	0.004
Trust	Trust in competent data reuse: pharmaceutical companies that produce medicines	No trust	2.81	<0.001
Trust	Trust in competent data reuse: the government	No trust	2.51	<0.001
Trust	Trust in competent data reuse: scientific researchers	Very little trust	2.22	<0.001
Trust	Provision of information regarding reuse of data	Via a website	2.14	<0.001
Trust	Trust in competent data reuse: scientific researchers	A little trust	2.01	<0.001
Trust	Trust in competent data reuse: the government	Very little trust	1.85	<0.001
Trust	Trust in competent data reuse: pharmaceutical companies that produce medicines	Very little trust	1.67	0.003
Trust	Trust in competent data reuse: the government	A little trust	1.42	0.006

Table 2. Statistically significant predictors for both no preference for, and don't know what kind of consent procedure (reference category opt-out)

Theme	Predictor	Category	OR	p-value	OR	p-value
			No preference		Don't know	
Trust	Provision of information regarding reuse of data	No permission to reuse	2.95	0.002	7.58	<0.001
Trust	Trust data reuse scientific researchers	Very little trust	3.40	<0.001	4.33	<0.001
Politics	If parliamentary elections were held today what is the percent chance that you will vote	0-25	2.05	0.039	2.94	0.013
Politics	Did you vote in the most recent parliamentary elections held on 17 March 2021	Not eligible to vote	3.40	0.001	4.56	<0.001
Politics	Are you very interesting in political topics fairly interested or not interested	Very interested	0.64	0.038	0.41	0.022
Media	Do you have the confidence that the following social media respect your privacy Facebook	Complete	2.18	0.008	2.35	0.031
Media	Do you have the confidence that the following social media respect your privacy Facebook	Not at all	0.38	<0.001	0.37	0.011
Health	This past month I felt so down that nothing could cheer me up	Seldom	1.62	0.015	1.89	0.029

4. Discussion

Insight in which patient characteristics are associated with preferences for consent procedures and the willingness to share health data are essential for a learning health system, and contribute to a better understanding of the representativeness of study populations in the Netherlands. It helps gain a deeper understanding of patients who choose not to share health data, helps researchers identify potential inclusion biases in their studies and aids to shape national governmental communication strategies regarding data sharing and EHDS.

Previous research shows that patients' preferences regarding healthcare decisions are dependent on their personal characteristics, such as demographic and socioeconomic factors. [12, 13] Our results are in line with these studies, as we see differences in socioeconomic factors and the preferences of the respondents. Studies show beliefs and attitudes also play a role in the preferences of patients and their decision making. [9, 10] Our results show specifically beliefs and attitudes, especially regarding media, politics and trust are correlated to the preferences of respondents. For example, respondents who preferred an opt-in procedure were more likely to have low levels of trust in scientific researchers or the government, compared to those who preferred an opt-out procedure. Notably, most respondents expressed a preference for an opt-in consent model, and the EHDS consent approach may not align with public preferences.

Due to the effects of the choice for a particular consent procedure and its relation to the reliability of study outcomes, minimizing these biases and why they exist is essential to ensuring high-quality scientific research. From a policy perspective, this knowledge is crucial as well, as it can inform evidence-based strategies to promote equitable participation and data sharing, while ensuring that policies are aligned with public attitudes and concerns. This is particularly relevant in the context of the European Health Data Space, where the proposed opt-out system may not align with the preferences of many citizens. Implementation of the EHDS without sufficient attention to these concerns risks creating an increasing group of dissatisfied individuals, potentially undermining trust and participation in health data initiatives.

5. Conclusions

Characteristics, beliefs and attitudes, especially regarding media, politics and the trust respondents have, are correlated to the preference for a specific consent procedure. Respondents who do not have a preference show similar characteristics to those who do not know. Population representativeness impacts data quality and, ultimately, the reliability of study outcomes. Minimizing these biases and why they exist is essential to ensuring high-quality scientific research and implementation of health data initiatives requires sufficient attention to these risks and the individuals behind routine health data.

References

[1] Enticott J, Johnson A, Teede H. Learning health systems using data to drive healthcare improvement and impact: a systematic review. BMC health services research. 2021;21:1-16. doi: 10.1186/s12913-021-06215-8

[2] Mc Grath-Lone L, Jay MA, Blackburn R, Gordon E, Zylbersztejn A, Wiljaars L, et al. What makes administrative data research-ready? : A systematic review and thematic analysis of published literature. International Journal of Population Data Science. 2022;7(1). doi: 10.23889/ijpds.v6i1.1718

[3] Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. Journal of clinical epidemiology. 2005;58(4):323-37. doi: 10.1016/j.jclinepi.2004.10.012

[4] Shabani M, Yilmaz S. Lawfulness in secondary use of health data: Interplay between three regulatory frameworks of GDPR, DGA & EHDS. Technology and Regulation. 2022;2022:128-34. doi: 10.71265/hawesm05

[5] Marcus JS, Martens B, Carugati C, Bucher A, Godlovitch I. The european health data space. IPOL| policy department for economic, scientific and quality of life policies, European Parliament Policy Department studies. 2022.

[6] de Man Y, Wieland-Jorna Y, Torensma B, de Wit K, Francke AL, Oosterveld-Vlug MG, et al. Opt-in and opt-out consent procedures for the reuse of routinely recorded health data in scientific research and their consequences for consent rate and consent bias: Systematic review. Journal of medical Internet research. 2023;25:e42131. doi: 10.2196/42131

[7] 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(5):e185. doi: 10.2196/jmir.9134

[8] Emam KE, Jonker E, Moher E, Arbuckle L. A review of evidence on consent bias in research. The American Journal of Bioethics. 2013;13(4):42-4. doi: 10.1080/15265161.2013.767958

[9] Jones RD, Krenz C, Gornick M, Griffith KA, Spence R, Bradbury AR, et al. Patient preferences regarding informed consent models for participation in a learning health care system for oncology. JCO Oncology Practice. 2020;16(9):e977-e90. doi: 10.1200/JOP.19.00300

[10] Russo S, Jongerius C, Faccio F, Pizzoli SF, Pinto CA, Veldwijk J, et al. Understanding patients' preferences: a systematic review of psychological instruments used in patients' preference and decision studies. Value in Health. 2019;22(4):491-501. doi: 10.1016/j.jval.2018.12.007

[11] Scherpenzeel AC. ""True" Longitudinal and Probability-Based Internet Panels: Evidence From the Netherlands. Social and behavioral research and the internet: Routledge; 2018. p. 77-104.

[12] Levinson W, Kao A, Kuby A, Thisted RA. Not all patients want to participate in decision making: a national study of public preferences. Journal of general internal medicine. 2005;20(6):531-5. doi: 10.1111/j.1525-1497.2005.04101

[13] Arpey NC, Gaglioti AH, Rosenbaum ME. How socioeconomic status affects patient perceptions of health care: a qualitative study. Journal of primary care & community health. 2017;8(3):169-75. doi: 10.1177/2150131917697439