# Erasmus School of Health Policy and Management

# Preferences for sharing health data digitally: a comparison between the Netherlands and Sweden

A thesis submitted to the Erasmus School of Health Policy and Management, Erasmus University Rotterdam for the partial fulfilment of a Master of Science degree in Health Economics, Policy, and Law

Name - Melaku Birhanu Alemu (536387ma@eur.nl)

Student Number - 536387

Supervisor – Dr Jorien Veldwijk (veldwijk@eshpm.eur.nl)

Word count - 9277

June 23, 2021

Rotterdam, Netherlands

# Abstract

**Background**: Digital health data is being hugely produced due to digitalisation in the health and other sectors. Different actors promote sharing of health data for disease surveillance, research, and policy improvement, which raised trust concerns on privacy issues. However, limited evidence exists on the general public's preference and heterogeneity for sharing of digital health data.

**Method**: A Bayesian D-efficient discrete choice experiment (DCE) design with 16 choice tasks containing five attributes (data collector, data shared with, the reason for data sharing, being informed, and review committee) was conducted as part of the "Research Towards Improving the Governance of Health Data in Cyberspace" research project. A descriptive and choice modelling analysis was performed. A mixed logit model and latent class logit model were used to account for individual and group level heterogeneity. Relative importance, acceptability rate and marginal change of acceptability were calculated and compared with countries.

**Result**: A total of 361 and 481 respondents from the Netherlands and Sweden completed the survey and were included in the analysis. On average, respondents do not prefer a technological company as a data collector and do not prefer to share their health data with such a company. Likewise, respondents do not prefer to share health data for marketing purpose. Lastly, participants prefer to be informed on data sharing and prefer a review committee to overseeing the data transfer. In addition, for displaying significant heterogeneity in both countries, three latent classes were identified. "Country", "gender", "use of application for health purpose", and "publication of health information on electronics" significantly explained class membership. The predicted acceptability rates ranged between 10% and 89% in the Netherlands and Sweden, depending on the situations. Similarly, the marginal increase in acceptability rate ranges from 0.6 to 27.9 percentage points in both countries, depending on the changed attribute level.

**Conclusion**: The willingness of the Dutch and Swedish general public to share health data digitally is influenced by the data collector, the company that the data shared with, the reason for data sharing, being informed and the review committee. There is individual and group-level heterogeneity for all the attributes. Therefore, it is recommended to consider the preference for the general public and country differences in designing data sharing platforms and regulations in the Netherlands and Sweden.

# Table of contents

# List of Tables

## Additional tables

# List of Figures

# Abbreviations and acronyms

| | |
|---|---|
| CLM | Conditional Logit Model |
| DCE | Discrete Choice Experiment |
| EHR | Electronic Health Record |
| GDPR | General Data Protection Regulation |
| LCM | Latent Class Model |
| MIXL | Mixed Logit Model |
| NHS | National Health Service |
| UK | United Kingdom |
| WHO | World Health Organization |

# Chapter 1: Introduction

Health data is being shared with different stakeholders for disease surveillance, education, research, and improving the health system (1). As the general public is the primary source and owner of health data (2), health data sharing should involve clinicians, researchers, and patients to work collaboratively to agree on data usage upfront and meet the patient's expectations (2, 3). However, most patients are unaware that their health data is being shared (4), which could raise ethical concerns on informed consent (5). In addition, some patients who do know about their health data sharing find the system unacceptable (4).

Digitalisation advancement creates substantial electronic data generation, storing, and analysing in the health sector and other technology companies (3, 6). Data sharing is supported and promoted by publishers (3), governments (7) and funding agencies (1); however, clinicians and researchers tend to be reluctant to share health data with other health organisations or companies. Academic research institutions devote time and resources to a health data collection that they want to use themselves first without publication competition (3). In addition, the strict ethical issues in health data sharing make academic institution hesitant to share health data, which may affect the timely use of data for decision making (3).

The World Health Organisation (WHO) urges countries to have a long-term strategic plan to develop and implement eHealth service to support health services, surveillance, literature, education, knowledge, and research (1). As a result, there is a remarkable growth in the national Electronic Health Record (EHR) system adoption in the past fifteen years. By 2016, more than half of the upper-middle and high-income countries adopted a national EHR system (8). The process of rushing to make digital health an integral part of countries' policies should be governed with transparency, accessibility, replicability, interoperability, privacy, security, and confidentiality data sharing principles (9). This could facilitate the adoption and expansion of digital technology through increasing public trust (9, 10).

The sharing of sensitive health data could happen accidentally due to coding errors or intentionally to develop new health technologies. For instance, the incidence of the National Health Service (NHS) data breach in England resulted in the sharing of confidential data of 150,000 patients in the three years period due to coding error (11). Besides accidental health data leaks, the United

Kingdom (UK) government plans to share sensitive health data with big technological companies to trace and develop effective medical technology (7). However, sharing of health data with technological companies raises concerns by the general public regarding privacy (12).

Pharmaceutical and academic research institutions are interested in patient's health data to develop safe and effective medical technologies and research. In their article on data sharing dilemmas, Jams and Toby (13) argued that health data is being shared with pharmaceutical companies without full permission from the participant. The main reasons for the compromised consent process arise from the difficulties of knowing the type of data before conducting clinical research. The other argument is a new data user could use the data after the study is completed (13). In addition to pharmaceutical companies, academic and research institutes are interested in health data for teaching and research. Even though it is less evident in Europe, compromised health data sharing principles by academic research institutes, hospitals, and healthcare providers are observed in different countries (14-16).

In recent years, technological companies engaged in accessing, collecting, and storing health and lifestyle information of the general population. In 2018, Apple company announced a partnership with thirteen major health care organisations in the US, including the University of Pennsylvania and Johns Hopkins, which allows the company to download patient health data into its device with consent from the patient (17). Besides collecting health data after patient consent, Apple and smartwatch/phone products automatically collect, store, and analyse data on people's lifestyle (step counting, food diaries, body weight, sleep rhythms, calorie use etc.) (18). This practice could be realised as an opportunity for patient empowerment. Still, it creates uncertainties on fraud, data abuse and patient privacy as the data is stored and analysed by a third party (17).

The Oviedo Convention on Human Rights and Biomedicine article 10 states that "Everyone has the right to respect for private life in relation to information about his or her health" and "Everyone is entitled to know any information collected about his or her health. However, the wishes of individuals not to be so informed shall be observed" (19). Therefore, the government has a positive obligation to respect, protect and fulfil the commitments in safeguarding patient privacy (20). In doing so, the general public's preference for data sharing needs to be considered in preparing rules and regulations that guide data sharing in cyberspace. Besides, preference studies in digital health

data sharing are essential for patients, clinicians, researchers, regulators, and policymakers by giving information on crucial aspects for the general public in data-sharing decisions.

However, as to the researcher's knowledge, very little is known about the general public preferences for sharing health data digitally. For example, the relative importance of attributes, the acceptability rate and marginal acceptability is unknown. In addition, the preferences of different populations in Europe might differ even though data sharing can happen across borders and easily in the EU due to the regulations (2). Therefore, this preference study is crucial to understand patient preference better to develop a governance system that can meet the general public's expectations and prevent data sharing breaches. In addition, it gives evidence to respond to present and future challenges that come with health sector digitalisation.

This research aimed to measure the general public preferences for health data sharing digitally and compare the two European Union countries (Netherlands and Sweden) to assess cross country preference heterogeneity. Both countries are similar in adopting the General Data Protection Regulation (GDPR) law (2), provide patient access to EHR systems (21, 22), and have national data protection authorities (13, 14). Therefore, the following research questions were formulated:

- What are the general public's preferences for digital health data sharing digitally in the Netherlands and Sweden?
- Do preferences for digital health data sharing differ between the Netherlands and Sweden?

The following general objective was formulated to answer the research questions.

- Assess the preferences for sharing health data digitally in the Netherlands and Sweden and compare the outcomes between the Netherlands and Sweden

To specifically explore the general objective of the study, the following specific objectives were framed:

- Assess the preferences for sharing health data digitally in the Netherlands and Sweden
- Quantify the heterogeneity of preferences in sharing health data digitally in the Netherlands and Sweden
- Estimate the relative importance of attributes for sharing health data digitally in the Netherlands and Sweden

3

- Calculate the acceptability rates and marginal change in acceptability rate for a particular sharing health data digitally situation in the Netherlands and Sweden
- Compare all the outcomes mentioned above between the Netherlands and Sweden

# Chapter 2: Theoretical Framework

## 2.1 Data sharing regulations and violations

The GDPR has six criteria for the processing of personal data, which includes (1) consent, (2) performance of a contract, (3) legitimate interest, (4) vital interest, (5) legal requirement, and (6) public interest (2). In 2019, the Dutch data protection authority announced a fining structure for the violation of the GDPR law (23).

However, violations in data sharing, storage, processing, and analysis were observed in different circumstances. In the UK, a company that worked on pregnancy and parenting packages sold user information of 34.4 million customers to other companies without consent (24). Similarly, the Norwegian data protection authority reported that Østfold HF Hospital stored personal data without sufficient protection (25).

Health data sharing with technological companies have been seen in recent years. For example, the NHS deals with technology companies like Google, Microsoft, Amazon, Faculty and Palantir for data sharing, which leads to over 13,000 petitions for publicity of the contract for privacy concerns (12). In addition, the UK government recently decided to share COVID-19 data, including sensitive patient information, with big tech companies to trace the pandemic (7).

Therefore, if we are able to know general pubic preferences for data sharing, we might have been able to predict the acceptance of people about GDPR violations. In addition, we could understand the acceptable data sharing options.

## 2.2 Preferences for data sharing and governance mechanisms

For effective health care program implementation, patient and general public preferences need to be considered to increase satisfaction and adherence to the intervention. The inclusion of the general public or patient preferences in policy decisions facilitates policy adoption (26, 27).

Although the general public has limited awareness of the existing practice of data use and sharing (28), there is a growing worldwide acceptability in sharing and linkage of sensitive health-related data electronically (10, 28, 29). The rising willingness to share health-related information digitally

resulted from amongst others trust in the security protocol for health data protection, confidence in the institution that uses the data and the perception that the data will be used to benefit the patients (28). However, the general public's acceptability for health data sharing is conditional on confidentiality, consent, public benefit, trust in the sharing company, and monitoring authority (28, 30). The data sharing and linkage conditions in the health sector might not be equally important for the population. For example, a study conducted in Scotland revealed that the type of data being shared (From general practitioner, from private sector, from education institution, from employer) is twice as important as the researcher type (University researcher, NHS, commercial companies for marketing, pharmaceutical companies) (31). Generally, patients are not comfortable with a permissive system (less formal consent, including sensitive information, broader users). Furthermore, preferences variation by sociodemographic characteristics for health data sharing is observed (32).

The contextual integrity theory developed by Helen Nissenbaum explains the theory of privacy in shaping the acceptability of sharing sensitive health-related information. The theory has five parameters in the data transfer option to be accounted for; (1) the data subject, (2) the data sender, (3) the data receiver, (4) information type and (5) transmission principle (33). The privacy of health data while sharing should be protected throughout the information flow (34). The theory of contextual integrity can be used to conceptualise the general public's preference and relative importance for the data transfer parameters.

## 2.3 Measuring preferences

A discrete choice experiment (DCE) an economical method to measure the stated preferences of an individual or a population and can be used to estimate the relative importance of various attributes in the analysis (35). In the past decade, there was an increasing use of DCE methods in healthcare to measure a stated preference of different population (26). It was introduced in healthcare as an economic evaluation framework to measure beyond health outcomes to include "non-health outcome" and "process attributes" in the healthcare system (35). The DCE can be used to estimate population preference and relative importance of attributes, trade-offs between attributes, acceptability rate and choice predictions (35, 36).

The DCE designs are consistent with economic theory and can predict real market situations. From an economic perspective, individuals make a combination of choices to maximise utilities (27). The choice experiment is founded on Lancaster's value theory which states that utility gets from the characteristics of the object (attribute) (37), and random utility theory, which states that utility has systematic and random components (38). Evidence showed that a DCE could accurately predict healthcare choices if it accounts for scale and preference heterogeneity (39).

The utility of the population cannot be observed directly, but we can estimate the value using observed choices (27). The people's utility for sharing the health data can be assessed using observed and unobserved characteristics. The indirect utility function using random utility theory is shown as:

$$U_i = V(\beta, X_i) + \varepsilon_i$$

Where; "V" represents the measurable utility; "$X_i$", vector for the attribute level; "$\beta$", vector for estimated coefficient and "$\varepsilon_i$", random error term (36).

The DCE can be conducted by presenting the respondent with a choice task with different characteristics (attributes) and levels. The experiment assumes that the attributes describe the factors which determine the decision for sharing health data information (35). A hypothetical situation can be used to elicit the respondents' stated preference with different levels of attributes (26) (see Figure 2 on section 3.6).

# Chapter 3: Research Methods

## 3.1 Project description

This study is a part of a collaborative research project between the University of Oxford and the University of Oslo, Uppsala University, and the University of Iceland. The project is entitled "Research Towards Improving the Governance of Health Data in Cyberspace" to explore the knowledge about electronic health data use and assess the general population's preference in different European countries (40).

The project has three main objectives: the first being exploring the citizens' expectations about data use and governance mechanisms and factors that affect their decision. Secondly, investigation of existing law, regulations and policies are capable of protecting health data. Finally, develop recommendations that ensure the use of health data in cyberspace that meets the population's expectations, security, and privacy.

## 3.2 Population characteristics

The data was collected in the Netherlands from April to May 2021. The Netherlands is located in Western Europe and part of the Caribbean. The Netherlands is the fives densely populated country in Europe (41), and the population density differences between regions with the highest 1374 people/ km2 at Zuid-Holland and the lowest 188 people/ km2 in Drenthe (42). According to the recent CBS report, the Netherlands has 17. 5 million estimated populations. Women's account for 50.6% of the population. On average, 53% of the population is aged 25 to 65 (42).

Swedish data was collected from March to April 2021. Sweden has an area of 450,295 square kilometres, which makes it the largest country in Northern Europe and the third-largest country in the EU. It has a population of 10.4 million inhabitants with a low population density (25people/ km2). Nearly half (51%) of the population is in the age range of 25-65 (43).

## 3.3 Study design

A societal perspective DCE was conducted in the Netherlands and Sweden. A cross-sectional ranking survey was conducted to select the best attributes in the Netherlands. For Sweden, the ranking survey is implicitly done as part of attribute development (see section 3.5.2).

## 3.4 Sampling

The samples were randomly selected using gender and age as a quota from the Netherlands and Sweden's general population. The samples were drawn from all permanent residents of both countries with ages above eighteen.

## 3.5 Data collection methods

### 3.5.1 Development of data collection tool and variables

A structured data collection tool was used to perform the DCE. The data collection was conducted through an online platform provided by a Survey engine vendor company in both countries (44). The survey was organised in three parts: sociodemographic, DCE, and subjective questions on the trust of various sectors (technological companies, the legal system, academic research institutions, and government).

### 3.5.2 Attribute development

The attribute development was conducted before this thesis project as part of the larger project in Sweden, Norway, UK and Iceland (40). The attributes and attribute levels were determined using three phases. The first step was attribute identification. A literature review and expert discussions were conducted in this phase to identify potential attributes and attribute levels for the DCE. The second step was attribute development. The development phase was mainly focused on further selecting the identified attributes valuable for the general public in making decisions for sharing health data digitally. A focus group discussion and ranking exercise with the target population was done to develop the attributes and levels. The final phase of attribute development was attribute refinement. As a final step, expert discussion and cognitive interviews were conducted to select the final attributes and attribute levels that need to be used for DCE (45).

For the Netherlands an additional ranking survey was conducted to make sure important factors were included in for the final DCE which determine decision making for sharing health data digitally. This attribute selection process is part of this thesis project. The survey was done as a way to confirm the attribute selection for the DCE that was done in Sweden. A total of 12 potential important factors when sharing health data was used to be ranked by Dutch general public.

Fifty-five respondents started the ranking survey, and 42 participants finished the ranking survey to confirm attributes for the final DCE survey. The age of respondents for the ranking survey was between 19-74, with a mean age of 50.1 years with a standard deviation of 16.3. Majority of the participants (52.4%) were males. The Dutch general public gave high ranking for the purpose of use of data and low ranking for health data sharing for profit making. Figure 1 shows the relative ranking result of suggested attributes for the Dutch general public. For detail description of potential attribute included for ranking survey see the Additional Table 1.
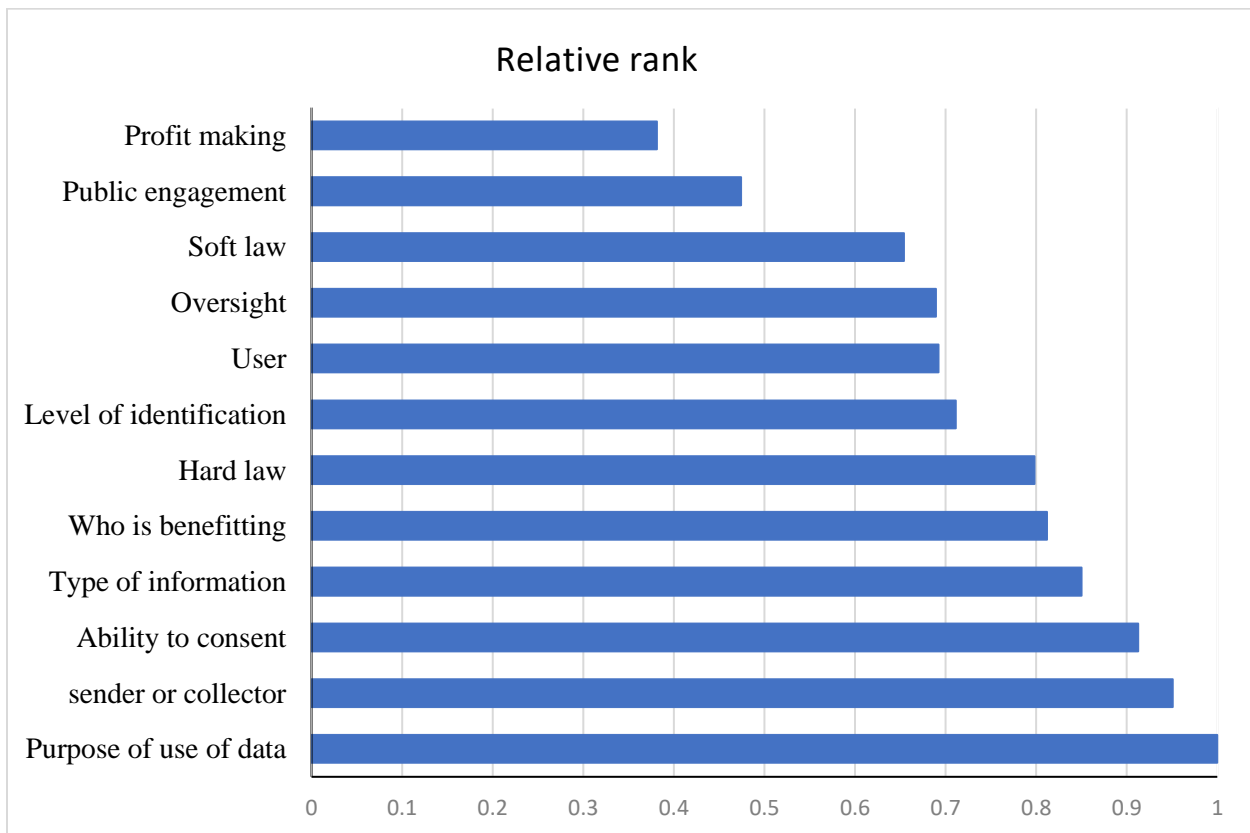


*Figure 1: Relative ranking of survey for attribute confirmation, Netherlands*

Based on the ranking survey and considering other attribute identification methods, five attributes were selected for the final DCE survey. Table 1 summarises the attributes used for the Dutch and

Swedish DCE survey. Data collector, the recipient of the data, reason for data sharing, information on data sharing and availability of review process were used as an attribute for this study.

*Table 1: Attributes and attribute level used in the discrete choice experiment to elicit the population preference for health data sharing digitally in the Netherlands and Sweden*

| Attribute | Level |
| --- | --- |
| **Data collector** | Technological companies |
| | General practitioners |
| | Academic institutions |
| **Data Shared with:** | Technological companies |
| | Pharmaceutical companies |
| | Authorities |
| | Academic institutions |
| **Reason for sharing** | Marketing purposes |
| | Drug development purposes |
| | Policy development purposes |
| | Quality control purposes |
| **Information on data sharing** | Not informed |
| | Informed with an opt-out option |
| | Only informed |
| | Informed with consent |
| **Availability of review process** | No review |
| | Review the transfer and use |
| | Review the transfer only |

## 3.6 Experimental design

The experimental design was developed before this thesis as part of the research project (40, 46). A Bayesian D-efficient design was created by Ngene software using 500 Halton draws and 1000 repetition. Expert guessing was used for constructing the pilot design as a prior information.

A pilot study was conducted with 28 unique choice tasks, which were arranged in four blocks randomly. Each respondent was randomly assigned to one of the four blocks to answer seven choice tasks. The blocking was used to increase the response efficiency by decreasing the cognitive efforts of the respondents who participated in the study (47). Interaction between attribute levels was assumed to be zero, and level balance was optimised. A total of 200 respondents from the UK, Norway, Iceland and Sweden participated in the pilot study (50 respondents each).

The final DCE design prior was updated from the conditional logit model results of the pilot study. Thirty-two unique choice tasks were generated in 4 blocks. Each block has one of the hypothetical health information types (physical health, mental health, lifestyle information, and genetic information). Respondents were randomly assigned to two blocks and asked to evaluate the choice tasks (sixteen choice task) associated with the blocks. The data sharing situations was presented and participants were asked to evaluate the unlabelled choice tasks whether the situation is acceptable for them or not. Figure 2 shows an example of a DCE task with a mental health information hypothetical situation.

Imagine that **mental health information (e.g. any diagnoses such as fatigue, depression, dementia or ADHD)** about you has been collected. The organisation that collected your data is now proposing to share it. Please read the situation description below and let us know if you think that sharing this information in this situation is acceptable or not.

(7 of 8)

| The organisation collecting my information is | my health care provider. |
| The reason they want to use my information is to | promoting, advertising or marketing their product or service. |
| They will share it with | a technological company. |
| When they share my information, I will | be offered information and an opt-out. |
| There will be | a committee that reviews the sharing of information. |

**Do you think this situation is acceptable?**

| ○ Yes | ○ No |

*Figure 2: The discrete choice experiment task to elicit the preference of the respondents on data sharing*

## 3.7 Data management and analysis

### 3.7.1 Data cleaning

Data cleaning and checking were conducted to check for incomplete responses and coding errors. Incomplete responses were discarded. Respondents who finished earlier than five minutes were excluded from the study as random answers could inflate the standard error as the choices made could be random.

### 3.7.2 Data analysis

The data were analysed using STATA version 16 software and excel. The data was presented using tables, figures and charts. The frequency, mean and standard deviation of the findings was presented descriptively. A significant proportion difference between the Netherlands and Sweden was checked for the respondents' characteristics using Chi-square. Categories that do not fulfil the chi-square assumption was dropped from the test.

A conditional logit model (CLM) was used to explore the data for preference and heterogeneity by using the country as interaction with all attributes. In addition, a subgroup analysis was conducted in the Netherlands and Sweden to further investigate the heterogeneity in the preference for sharing health data. As choices are made at the individual level and policies are made at an aggregate level, it is crucial to investigate the individual and group level heterogeneity in the preference for sharing health data digitally. The Mixed Logit Model (MIXL) and Latent Class Model (LCM) were used to understand individual and group-level heterogeneity in the Netherlands and Sweden.

A MIXL was fitted for the Netherlands and Swedish dataset separately to better compare the acceptability rate of health data sharing digitally between the Netherlands and Sweden. The model uses simulated maximum likelihood methods to estimate the individual level and population parameters by considering choices for health data sharing as a dependent variable and attributes as the independent variable (36, 48). The model was estimated using 1000 Halton draws to estimate the average and individual level coefficients for the Netherlands and Sweden. All the attributes are included as a random variable to see preference heterogeneity. Based on the individual level utility score, the relative importance of attribute and acceptability rate for digital health data sharing was calculated using different scenarios.

The utility was estimated using the following model:

$U_{yes} = \beta_{techcollect} \times$ Technological company $+ \beta_{academiccollect} \times$ Academic Research Project $+ \beta_{techshare} \times$ Technological company $+ \beta_{authorityshare} \times$ National Authority $+ \beta_{pharmashare} \times$ Pharmaceutical company $+ \beta_{policyinitiative} \times$ Policy Initiative $+ \beta_{newproductdevelopment} \times$ New Product Development $+ \beta_{promotion\&marketing} \times$ Promotion and Marketing $+ \beta_{consent} \times$ Consent $+ \beta_{review} \times$ Review Committee

Uopt-out=0

The relative importance of attributes was determined by subtracting the highest attribute level values from the lowest estimate in each latent class. The attribute with the largest differences value resembles the most important attribute and received a total score of 1. Their difference value was divided by the difference value of the most important attribute for all other attributes.

The acceptability rate was calculated using the results from the MIXL by using excel. First, the individual level acceptability rate was estimated based on individual level utility estimated using MIXL then the average acceptability rate of the individual rates was taken to compare the Netherlands and Sweden.

$$Prob(Y = j) = \frac{\exp{(V_j)}}{\sum_{k-1}^{J} \exp{(V_k)}}$$

The structural part of individual level utility is represented by V. Thus, the probability of choosing to share health data digitally, J, over the opt-out is calculated by dividing the exponent of the utility of the accepting to share health data by the sum of the utility of accepting the data sharing and the opt-out, K (36).

Finally, the marginal acceptability rate was calculated by making a difference in acceptability rate in changing one attribute from the base case. This was done by keeping other attributes constant while changing one attribute level from the base case.

To investigate the preference variation to identify group of respondents with similar preferences, the LCM was fitted on the pooled dataset by merging the Dutch and Swedish dataset. This method helps to evaluate the general public's willingness to trade-off attributes in different latent classes (26). The model assumes attributes have a heterogeneous effect and identifies unobserved classes in the response. In each class, the preference weight was estimated using conditional logit regression. The best-fitting model (appropriate number of classes) was determined using Akaike information criteria (AIC) and log-likelihood (36, 49). Sociodemographic variables, type of information, use of electronics for the purpose of health, and participation in research were tested for class membership to explain class membership for the best fitted LCM. All attributes were included as a dummy coded variable.

## 3.8 Data quality

The sampling was conducted using age and gender as quota criteria, enabling this research to have representative data of the general public. In addition to removing incomplete respondents from the analysis, individuals who completed the survey in less than five minutes was also excluded from the research to ensure data quality. Moreover, the data was collected electronically, which avoided missing value, data entry errors.

## 3.9 Ethical considerations

Ethical clearance was obtained from Sweden (*Dnr 2020-00623*) and ESHPM (*Ref. 21-011*). The purpose and importance of the study were stated in the questionnaires. Informed consent was obtained from each study participant. Participants were informed that they have the right to withdraw from the study if they face any inconvenience during the data collection.

## 3.10 Research dissemination

A final draft of the research document will be disseminated to the Erasmus University Rotterdam library. In addition, the finding will be presented to the Erasmus School of Health Policy and Management. Finally, the results will be published in an international journal in consultation with the supervisor.

# Chapter 4: Results

## 4.1 Respondent's characteristics

### 4.1.1 Sociodemographic information

A total of 408 and 510 respondents completed the survey in the Netherlands and Sweden, respectively. However, forty-seven respondents from the Netherlands and twenty-nine respondents from Sweden were excluded from the analysis as they completed the survey in less than five minutes.

*Table 2: Sociodemographic information of participants of health data sharing digitally discrete choice experiment, Netherlands and Sweden*

| Variable | Category | Netherlands | | Sweden | |
|---|---|---|---|---|---|
| | | Number (n=361) | Percentage (%) | Number (n=481) | Percentage (%) |
| Age ** | 18-35 (young adulthood) | 77 | 21.3 | 124 | 25.8 |
| | 35-55 (middle adulthood) | 126 | 34.9 | 145 | 30.2 |
| | 56 (older adulthood) | 158 | 43.8 | 212 | 44.1 |
| Gender | Female | 185 | 51.3 | 248 | 51.6 |
| | Male | 176 | 48.8 | 231 | 48.0 |
| | Other | 0 | 0.0 | 2 | 0.4 |
| Education** | High school or lower | 5 | 1.39 | 38 | 7.9 |
| | High School | 75 | 20.8 | 167 | 34.7 |
| | Professional qualification | 124 | 34.4 | 90 | 18.7 |
| | Bachelor's degree | 116 | 32.1 | 150 | 31.2 |
| | Doctoral degree | 40 | 11.1 | 18 | 3.7 |
| | Other | 0 | 0.0 | 18 | 3.7 |
| | Prefer not to say | 1 | 0.3 | 0 | 0.0 |
| Nationality** | The Netherlands | 355 | 98.3 | - | - |
| | Swedish | - | - | 445 | 92.5 |
| | Other | 6 | 1.7 | 36 | 7.5 |
| Job in the Health Sector** | Yes | 63 | 17.5 | 119 | 24.7 |
| | No | 297 | 82.3 | 359 | 74.6 |
| | Prefer not to say | 1 | 0.3 | 3 | 0.6 |

*\*\* Chi-square significant at 0.01 level; \* Chi-square significant at 0.05 level*

Nearly half of the respondents are female in both countries. The majority of the respondents, 158 (43.8%) and 212 (44.1%), are above 56 years old in the Netherlands and Sweden. Most Dutch respondents have a professional qualification, 124 (34.4%), followed by a bachelor's degree 116 (32.1%). However, utmost respondents in Sweden have a high school education of 167 (34.7%), followed by a 150 (31.2) bachelor's degree. There is no significant difference between male and female in both countries. However, a significant difference in respondents characteristics was observed in the Netherlands and Sweden. Table 2 shows the sociodemographic characteristics of participants from the Netherlands and Swedish survey.

Nearly 20% of respondents in the Netherlands have been a part of a research study. However, twelve per cent of respondents are not sure whether they have been a part of a research or not. Similarly, 17% of respondents in Sweden participated in a research study, and 14% of respondents do not know their participation in a research study.

## 4.1.2 Electronic technology use for health purpose

Majority of Swedish respondents 356 (74%) uses electronic health record. The use of electronic health record in the Netherlands is beneath half 160 (44.3%). The use of applications to monitor lifestyle (pedometer, heart rate, nutrition) is almost similar between the Netherlands (57%) and Sweden (58%). On average, 20% and 17% of respondents participated in a research study in the Netherlands and Sweden. Table 3 presents the electronic use practice of respondents in the Netherlands and Sweden.

*Table 3: Application use for health and lifestyle in the Netherlands and Sweden*

| Variable | Category | Netherlands | | Sweden | |
|---|---|---|---|---|---|
| | | Number (n=361) | Percentage (%) | Number (n=481) | Percentage (%) |
| Use of electronic health record** | Yes | 160 | 44.0 | 356 | 74.0 |
| | No | 181 | 50.1 | 113 | 23.5 |
| | Don't know | 20 | 5.5 | 10 | 2.1 |
| | Prefer not to say | 0 | 0.0 | 2 | 0.4 |
| Use of app use for (pedometer, heart rate, nutrition, etc.) * | Yes | 206 | 57.1 | 279 | 58.0 |
| | No | 154 | 42.7 | 197 | 41.0 |
| | Don't know | 1 | 0.3 | 3 | 0.6 |
| | Prefer not to say | 0 | 0.0 | 2 | 0.4 |
| | Every hour | 9 | 2.5 | 14 | 2.9 |
| | Multiple times a day | 24 | 6.7 | 33 | 6.9 |
| | A few times a day | 25 | 6.9 | 20 | 4.2 |
| Frequency of publishing health information in app** | Several times a week | 26 | 7.2 | 28 | 5.8 |
| | A few times a week | 18 | 5.0 | 26 | 5.4 |
| | Several times a month | 17 | 4.7 | 6 | 1.3 |
| | A few times a month | 12 | 3.3 | 22 | 4.6 |
| | Rarely | 59 | 16.3 | 99 | 20.6 |
| | Never | 154 | 42.7 | 193 | 40.1 |
| | I do not know | 16 | 4.4 | 39 | 8.1 |
| | Prefer not to say | 1 | 0.3 | 1 | 0.2 |

*App: Application; electronic health record: Vårdguiden's e-services for Sweden; ** Chi-square significant at 0.01 level; * Chi-square significant at 0.05 level*

## 4.2 Preference and heterogeneity exploration

The CLM with country as an interaction on the pooled dataset showed respondents' have a significant preference for all attributes. However, the preferences are significantly differing between the Netherlands and Sweden for all the attributes. For example, respondents do not prefer a technological company as a data collector and do not prefer to share with technological

companies. In addition, sharing health data for marketing does not prefer. Finally, respondents prefer to have informed on data sharing and the availability of the review committee (see Table 4).

*Table 4: Conditional logit model with a country as an interaction*

| Attribute | Level | Main estimate | | Country Interaction | |
|---|---|---|---|---|---|
| | | Coef. | Std. Err | Coef. | Std. Err |
| | Tech company | Ref | | | |
| Data collector | General practitioner | 0.33** | 0.06 | 0.24* | 0.10 |
| | Academic institution | 0.34** | 0.05 | -0.03 | 0.08 |
| | Tech company | Ref | | | |
| Data Shared with: | Pharma company | 0.34** | 0.05 | -0.21** | 0.08 |
| | Authorities | 0.48** | 0.06 | -0.45** | 0.09 |
| | Academic institution | 0.43** | 0.07 | -0.29** | 0.10 |
| | Marketing | Ref | | | |
| Reason for sharing | Drug development | 0.83** | 0.08 | 0.02 | 0.13 |
| | Policy development | 0.75** | 0.09 | 0.33* | 0.14 |
| | Quality control | 0.92** | 0.09 | 0.07 | 0.13 |
| | Not informed | Ref | | | |
| Information | Informed with optout | 1.27** | 0.10 | -0.33* | 0.14 |
| | Only informed | 0.92** | 0.08 | 0.00 | 0.12 |
| | Informed with consent | 1.30** | 0.10 | -0.06 | 0.16 |
| | No review | Ref | | | |
| Review | Review the transfer only | 0.81** | 0.08 | -0.53** | 0.11 |
| | Review the transfer and use | 0.82** | 0.08 | -0.52** | 0.10 |
| Optout | | 1.95** | 0.16 | -0.03 | 0.24 |

*\*\*P-value <0.01; \*P-value <0.05: Coef., Coefficient; Std. Err, Standard error; Ref, Reference category: Note: the dummy variable for country is (Netherlands=1).*

However, preferences differ between the Netherlands and Sweden. For example, the Dutch respondents have a significantly stronger preference for the general practitioner as a data collector than Swedish respondents. Similarly, the Dutch respondents have a significantly lower preference for pharmaceutical companies, authority, and academic institutions to share their health

information with than Swedish respondents (see Table 4). For individual and group level subgroup analysis for preference heterogeneity, see section 4.3 and 4.4.

## 4.3 Preferences for the Netherlands and Sweden: MIXL

The significant interaction of the country with all the attributes motivated the uses of the MIXL to understand preference difference between the Dutch and Swedish general public. The MIXL was conducted in a separate dataset between both countries. The Dutch and Swedish general public have individual-level heterogeneity in preferences for all the attributes, as shown in Table 5.

On average, Dutch and Swedish people prefer general practitioner or academic institution as data collector compared with technological companies. Similarly, Dutch people prefer to share health data with pharmaceutical companies compared to sharing it with technological companies. However, the Dutch respondents have no preference for sharing with authorities or academic institutions over technological companies. On the contrary, the Swedish prefers to share health data with authorities, academic institutions, and pharmaceutical companies compared with technological companies.

*Table 5: MIXL result of the Netherlands and Sweden*

| Attribute | Level | Netherlands | | | | Sweden | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Coef. | Std. Err | SD | Std.Err*SD | Coef. | Std.Err | SD | Std.Err*SD |
| Data collector | Tech company | Ref | | | | | | | |
| | General practitioner | 1.29** | 0.14 | 1.13** | 0.16 | 0.75** | 0.11 | 0.97** | 0.16 |
| | Academic institution | 0.71** | 0.11 | 0.18 | 0.30 | 0.70** | 0.10 | 0.02 | 0.35 |
| Data Shared with: | Tech company | Ref | | | | | | | |
| | Pharma company | 0.25* | 0.12 | 0.01 | 0.43 | 0.72** | 0.11 | 0.19 | 0.46 |
| | Authorities | 0.00 | 0.14 | 0.96** | 0.19 | 0.99** | 0.13 | 0.67** | 0.19 |
| | Academic institution | 0.12 | 0.14 | 0.60** | 0.21 | 0.88** | 0.13 | 0.52 | 0.33 |
| Reason for sharing | Marketing | Ref | | | | | | | |
| | Drug development | 1.75** | 0.16 | 0.85** | 0.18 | 1.81** | 0.14 | 0.87** | 0.19 |
| | Policy development | 2.19** | 0.17 | 0.81** | 0.20 | 1.62** | 0.14 | 0.87** | 0.18 |
| | Quality control | 2.04** | 0.16 | 0.32 | 0.25 | 1.97** | 0.15 | 0.60** | 0.21 |
| Information | Not informed | Ref | | | | | | | |
| | Informed with optout | 1.98** | 0.19 | 1.64** | 0.18 | 2.68** | 0.17 | 1.22** | 0.16 |
| | Only informed | 1.92** | 0.14 | 0.54 | 0.29 | 1.94** | 0.14 | 0.88** | 0.22 |
| | Informed with consent | 2.57** | 0.21 | 1.90** | 0.19 | 2.84** | 0.19 | 1.79** | 0.18 |
| Review | No review | Ref | | | | | | | |
| | Review the transfer only | 0.47** | 0.12 | 0.15 | 0.24 | 1.71** | 0.12 | 0.50* | 0.22 |
| | Review the transfer and use | 0.66** | 0.12 | 0.75** | 0.19 | 1.73** | 0.12 | 0.70** | 0.17 |
| Optout | | 4.03** | 0.31 | 2.64** | 0.17 | 4.04** | 0.28 | 2.75** | 0.15 |

*\*\*P-value <0.01; \*P-value <0.05: Coef., Coefficient; Std. Err, Standard error; Ref, Reference category: Std.Err\*SD, standard error of the standard deviation;*

*MIXL, Mixed Logit Model*

### 4.3.1 Relative importance: MIXL

When health data is being shared, being informed on health data sharing has high relative importance in both countries, which is followed by the "reason for sharing". However, compared with the Dutch respondents, Swedish respondents have a relatively high preference for the company that the data is being shared with and the availability of a review committee during data sharing. Similarly, the Dutch populations gave relatively higher importance for the "data collector" and "reason for sharing" compared with the Swedish respondents. "Data shared with" and "Data collector" has relatively low importance in choosing to share health data in both the Dutch and Swedish general public (see figure 3).
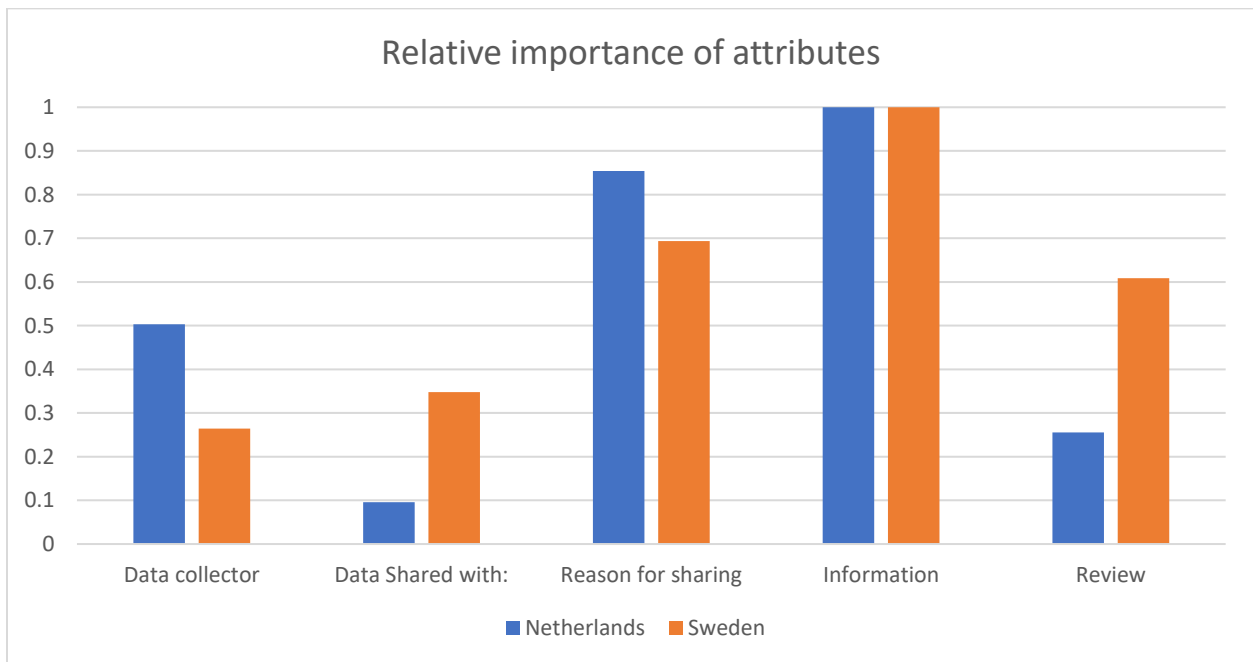


*Figure 3: Relative importance of attributes for the Netherlands and Sweden for sharing health data digitally*

### 4.3.2 Acceptability rate: MIXL

The acceptability of health data sharing digitally with the base case (technology company collects data, shared with the technological company, used for the purposes of marketing, no information was given and without review) is nearly 10% for the Netherlands and Sweden. The acceptability of respondents with the combination of the average best case in the Netherlands and Sweden is 80.5% and 89%, respectively (see Table 6).

*Table 6: Acceptability rate for health data sharing digitally in the Netherlands and Sweden*

| Scenarios | Netherlands | Sweden |
|---|---|---|
| Base case (worst case) | 10.3 | 10.0 |
| Best case based on Dutch MIXL | 80.5 | 85.3 |
| Best case based on Swedish MIXL | 76.9 | 89.0 |
| Base case + Informed consent | 35.76 | 38 |

*Note: the level change is from the base case keeping other attributes constant. **Base case**; Data collected by and shared with the technological company for the purpose of marketing without information and review committee. **Dutch best case**: is a general practitioner data collector, data shared with a pharmaceutical company for the purpose of policymaking with informed consent and review committee in the transfer and review phase. **Swedish best case**: general practitioner data collector, share with authorities for the purpose of quality improvement with informed consent and review committee in the transfer and use of the data*

### 4.3.3 Marginal effect analysis: MIXL

Informed consent has a high marginal effect on the acceptability of data sharing digitally in both countries. Providing informed consent could increase the acceptability of data sharing by 25.4 and 27.9 percentage points (PP) in the Netherlands and Sweden. Netherlands has a lowest marginal effect (0.6 PP) on acceptability for the change in data shared with from technological company to authority. Figure 4 shows the marginal change of attribute level from the base case. Note: the base case is Data collected by and shared with the technological company for the purpose of marketing without information and review committee.
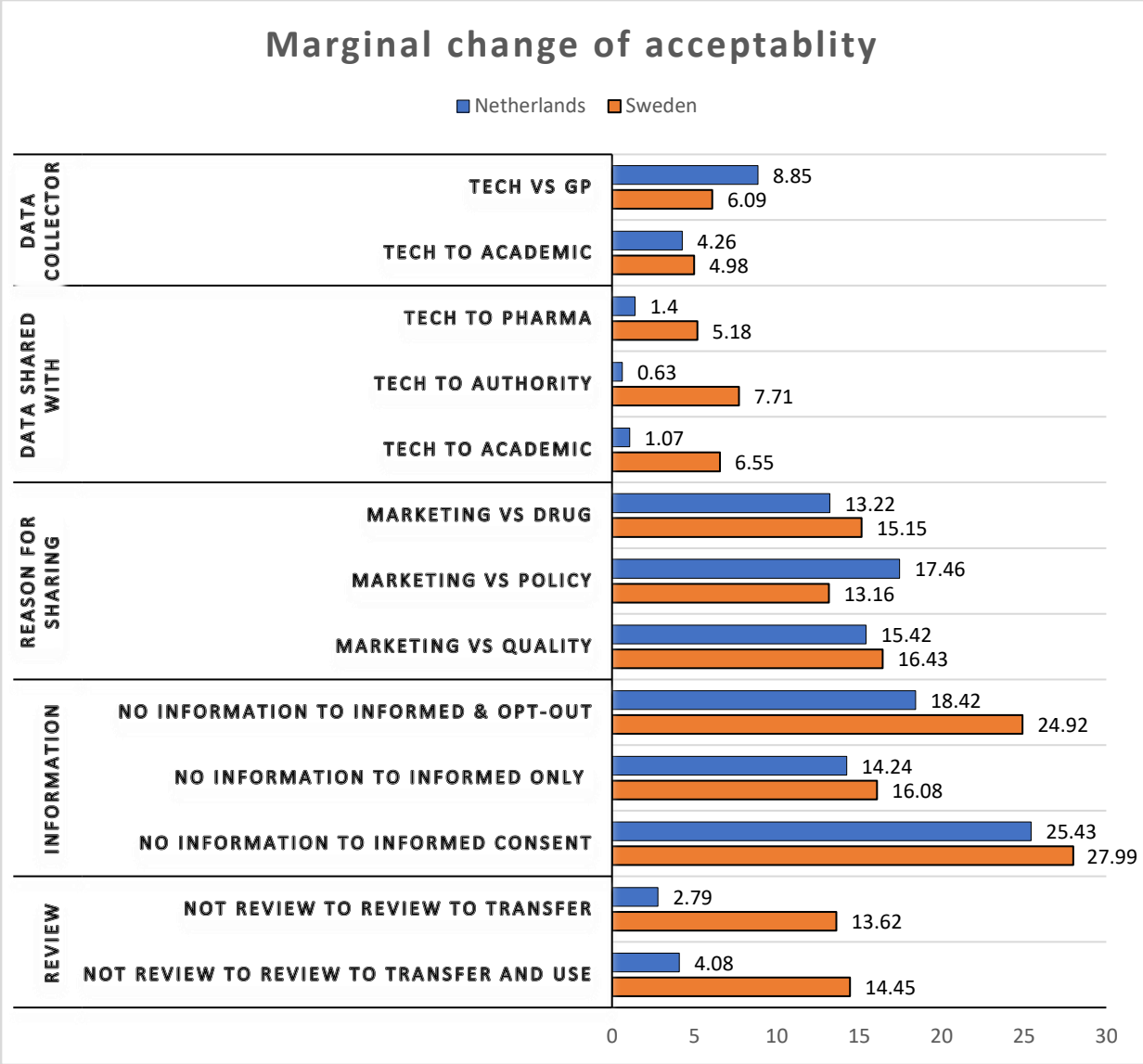
*Figure 4: Marginal change of attribute level from the base case between the Netherlands and Sweden.*

## 4.4 Pooled preferences and relative importance: LCM

### 4.4.1 Pooled preferences: LCM

A three-class solution model best fits the data for the Netherlands and Sweden based on the log-likelihood, Akaike information criteria (AIC) and Bayesian information criterion (BIC) (50), and sound in- perpetration of classes (51). Figure 5 shows the information criterion for different classes.
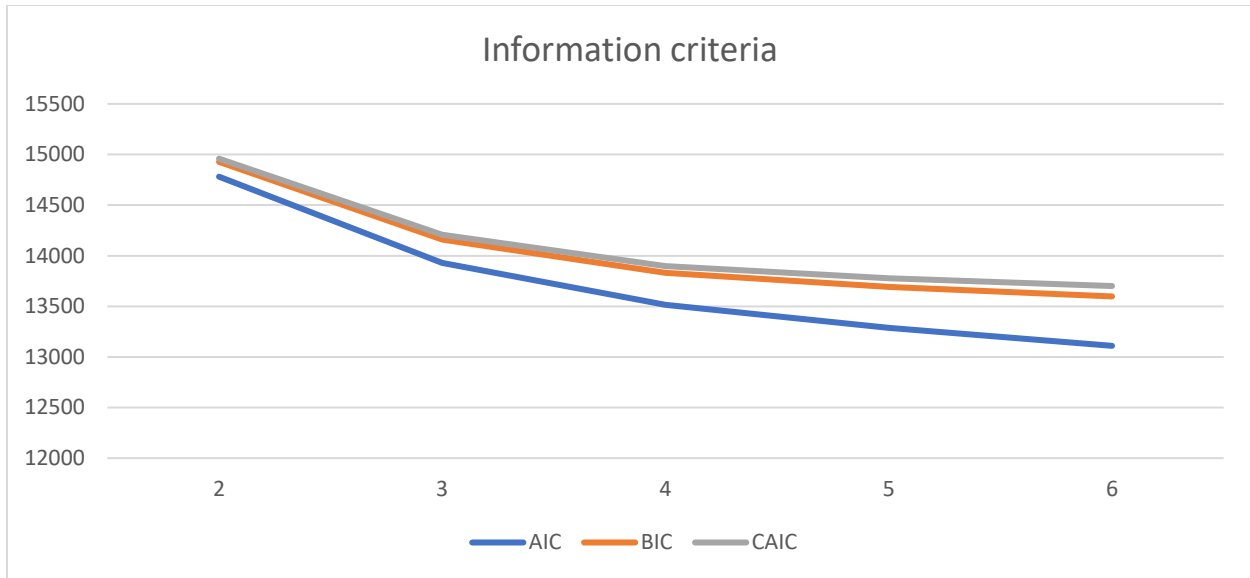
*Figure 5: Information criterion to determine the best fitting model*

The mean maximum probability of class assignment is 94.8%. Country, gender, application use for health-related activity and publishing health information frequently significantly determined individual membership in one of the latent classes. Dutch respondents are more likely to be in class 1 and 2 than class 3. However, male respondents are less likely to be in class 1 and 2 compare to class 3. People who use lifestyle application are less likely to be in class 1 relative to class 3. Respondents who publish health information frequently are less likely to be in class 1 and 2 compared to class 3. Table 7 presents the class membership variables.

Respondents from classes 1 and 2 prefer a general practitioner or academic institution over technological companies as health data collectors. Respondents have no statistically significant preference for general practitioner over an academic institution. Respondents in class 3 have no preference in data collector institutions.

Like the data collector, respondents in class 1 and 2 do not prefer to share their health data with technological companies compared to pharmaceutical authorities or academic institutions. Class 3 respondents have no preference for sharing health data between technology companies and other pharmaceutical, academic or authorities.

*Table 7: LCM with class membership variables in the Netherlands and Sweden*

| Attribute | Level | Class 1 | | Class 2 | | Class 3 | |
|---|---|---|---|---|---|---|---|
| | | Coef. | Std. Err | Coef. | Std.Err | Coef. | Std.Err |
| | Tech company | Ref | | | | | |
| Data collector | General practitioner | 1.42** | 0.17 | 0.56** | 0.08 | -0.08 | 0.28 |
| | Academic institution | 0.88** | 0.16 | 0.48** | 0.07 | 0.19 | 0.29 |
| Data Shared | Tech company | Ref | | | | | |
| with: | Pharma company | 0.39* | 0.18 | 0.37** | 0.08 | 0.26 | 0.32 |
| | Authorities | 0.66** | 0.18 | 0.32** | 0.08 | 0.74 | 0.42 |
| | Academic institution | 0.99** | 0.17 | 0.31** | 0.08 | 0.21 | 0.34 |
| Reason for | Marketing | Ref | | | | | |
| sharing | Drug development | 1.89** | 0.25 | 1.15** | 0.09 | 1.64** | 0.45 |
| | Policy development | 2.04** | 0.24 | 1.23** | 0.09 | 1.10** | 0.37 |
| | Quality control | 2.06** | 0.25 | 1.41** | 0.10 | 1.14** | 0.42 |
| | Not informed | Ref | | | | | |
| Information | Informed with optout | 0.88** | 0.19 | 1.87** | 0.10 | 1.60** | 0.44 |
| | Only informed | 0.79** | 0.15 | 1.45** | 0.09 | 1.80** | 0.42 |
| | Informed with consent | 1.09** | 0.23 | 2.15** | 0.11 | 1.29** | 0.41 |
| | No review | Ref | | | | | |
| Review | Review (transfer only) | 0.85** | 0.16 | 0.87** | 0.07 | 0.95** | 0.33 |
| | Review (transfer and use) | 0.68** | 0.15 | 0.97** | 0.08 | 1.02* | 0.43 |
| Optout | | 5.80** | 0.40 | 3.06** | 0.17 | -0.30 | 0.59 |
| Class share | | 22.8% | | 51.2% | | 26.1% | |
| Class membership | | | | | | | |
| Variable | | Share 1 | | Share 2 | | Share 3 | |
| Country (Netherlands) | | 1.29** | 0.23 | 0.72** | 0.20 | | |
| Gender (male) | | -0.57** | 0.22 | -0.80** | 0.18 | | |
| Use_app_health_(yes) | | -0.63* | 0.27 | 0.05 | 0.24 | Ref | |
| Publish_frequently (yes) | | -0.88** | 0.29 | -0.48* | 0.22 | | |
| Constant | | 0.25 | 0.24 | 0.99** | 0.20 | | |

*\*\*P-value <0.01; \*P-value <0.05: Coef., Coefficient; Std. Err, Standard error; Ref, Reference category*

All classes prefer sharing their data for dug development, policy development, and quality control than sharing for marketing purposes. Similarly, all classes desire to have informed consent or informed with opt-out or only informed rather than not being told about data sharing. Finally, all classes prefer a review committee that oversees data transfer or the transfer and use of data compared with sharing without a review process. Table 7 shows the three-class latent model.

## 4.4.2 Pooled relative importance of the attributes: LCM

There is group-level heterogeneity in the preference for sharing health data digitally (see figure 4). Reason for sharing health data is the most important attribute in class 1. Class 2 and 3 relatively prefer being informed on data sharing as an important attribute for health data sharing followed by the reasons for sharing health data in both classes. The availability of the review committee, the institute that the data shared with, and the data collector are the least important attributes in class 1, 2 and 3, respectively (see figure 6).
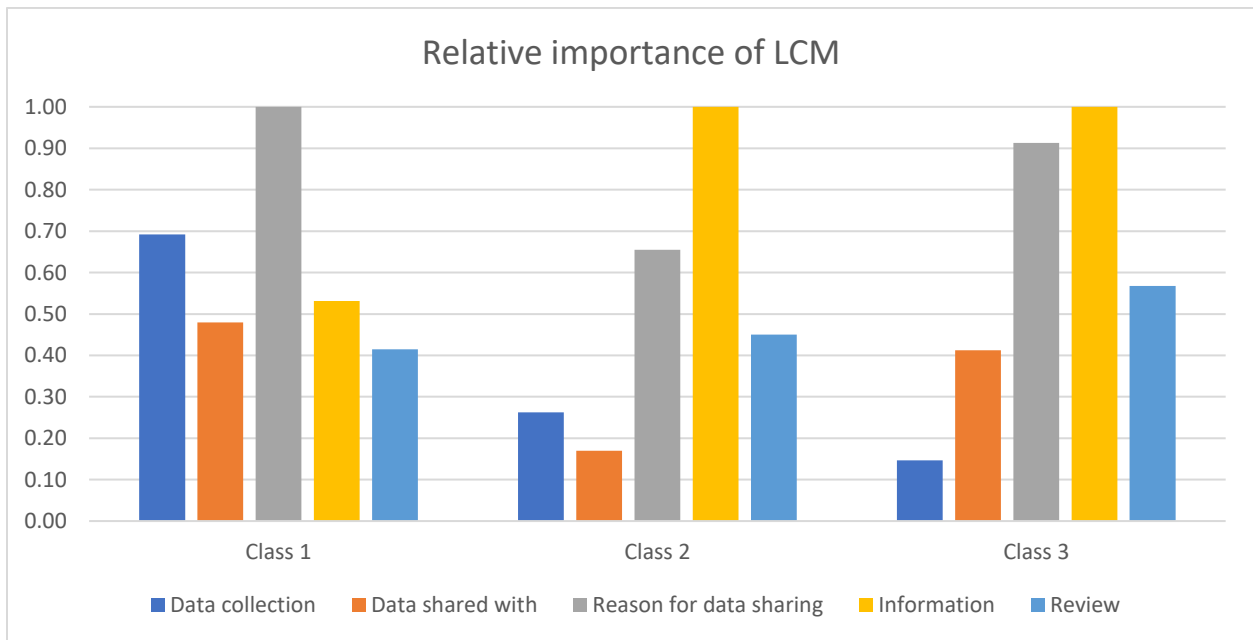


*Figure 6: Relative importance of attributes in different latent classes*

# Chapter 5: Discussion and Conclusion

This study revealed that the willingness of the Dutch and Swedish general public to share health data digitally is influenced by the health data collector, the company that the data is being shared with, the reason for data sharing, being informed and the existence of a review committee. In addition, three latent classes were identified in the pooled dataset of the Netherlands and Sweden. However, individual preference heterogeneity and aggregate level heterogeneity existed for sharing health data digitally in both countries. The preferences and heterogeneities for sharing digital health data with the relative importance, acceptability rate and marginal change of the attribute level are discussed below.

On average, the Dutch and Swedish respondents do not prefer a technological company as a health data collector compared to the general practitioner or academic institution. The result is similar to a study conducted in four European countries (Sweden, Norway, UK and Iceland on the general public's preference for sharing health data (46). The reason could be a lack of trust in technology companies due to frequent allegations on the GDPR violations in data privacy (52).  This is also in accordance with the Helen Nissenbaum theory of contextual integrity, stating that health data privacy should follow the information path, including data sender and receiver (33, 34). Therefore, the technology companies are advised to work closely with the academic institution in health data collection to increase the acceptability by the general public for digital health data sharing.

The Swedish respondents strongly prefer sharing health data with a pharmaceutical company, authority or academic institutions rather than technological companies. The finding aligns with a similar study conducted in Sweden, Norway, UK and Iceland, which prefers a pharmaceutical company over a technological company (46). The unwillingness to share health information with technological companies could be because of privacy concerns with technological companies (52). However, the Dutch respondents prefer to share health data with pharmaceutical companies but have no preference for authority or academic institutions over the technological companies. Nonetheless, a significant individual preference variation exists in sharing health data with the authority and academic institutions. Thus, the Dutch general public preference difference could arise from the individual-level significant heterogeneity for the authority and academic institutions.

The Dutch and Swedish general public prefers data sharing for drug development, policy preparation, or quality control than sharing for marketing purposes. The finding is supported by a DCE conducted in Sweden and Scotland that showed a high acceptability rate for data linkage for research purposes (53). In addition, the result is in line with a similar study conducted in Norway, Sweden, UK and Iceland, that respondents do not want to share their health information for marketing purposes (46). The GDPR, Article 6 support the provision of clear information on data sharing, stating that lawful data processing should involve consent for use for specific purposes (2). Therefore, the purposes of health data sharing should be explicitly mentioned to get informed consent.

The Dutch and Swedish general public do not prefer digital health data sharing without being informed beforehand over not being informed when data is being shared. The GDPR supports informed consent to ensure privacy and transparency as a fundamental principle on data sharing (2). Furthermore, informed consent is in accordance with international laws. The Oviedo Convention, Article 5 declares that any intervention should be conducted after a person gives free and informed consent (5). Therefore, in addition to the obligations by international law, the provision of informed consent could increase in data sharing acceptability for sharing digital health data.

Finally, the Dutch and Swedish general public prefer to have a review committee on the transfer of data or the transfer and use of health data rather than transferring without a review committee. The preference for a review committee's existence to oversee the health data sharing process could be supported by the general public desire for a conservative system than a permissive system (28, 30, 31). Therefore, the review committee should be available to meet the general public expectation and increase acceptability in digital health data sharing.

The relative importance of attributes is different for the Dutch and Swedish general public. In addition, the three latent classes have different relative importance for the attributes, meaning there are groups in the population that have a different preference for the attributes.

Being informed is the most important attribute for both countries. The result in line with a study on the preferences of the Norway and UK general public, which gives being informed the most important attribute (46). In addition, it is the most important attributes in the latent class 1 and 2.

This is in lines with the majority of sub-group analysis results in a similar study (46). Provision of informed consent before intervention is supported by the international conventions (Oviedo Convention Article 5) (19), and GDPR (Article 6(a) and Article 9(2a)) (2). Therefore, respecting the laws on informed consent could increase the acceptability of the Dutch and Swedish general public for sharing health data digitally, as it is the most important attribute for decision making.

The reason for data sharing is the second most important attribute next to being informed when data is shared for both countries. This result is similar to the Norway and UK general public preference for sharing health data digitally (12, 46). In addition, the reason for data sharing is the most important attribute in latent class 1 and the second most important attribute in the other latent classes. Like informed consent, the reason for data sharing is supported by the international law (Oviedo Convention Article 5), which states the participant must have information on the purpose of the interventions (19). Therefore, detailed information on the reason for the health data sharing should be provided in designing health data sharing arrangements to meet the general public's expectations and ensure informed decision-making. The detailed information on the reason and other aspects of data sharing would resolve partial or incomplete informed consent (13).

The relative importance of a review committee differs highly between different countries. It is the third and fourth most important attribute in the Sweden and Netherlands, respectively. In a similar study, the availability of a review committee is the third most important attribute for the UK general public (46). However, it is the most important attribute for the Iceland general public and the least important attribute for Norway (46)

Finally, the data collector is the third most important attribute for the Dutch people with the least being the company that the data is being shared. The Swedish general public have relatively low preference for data collector company. The result is in line with to a study conducted in Norway, Sweden, UK and Iceland (46). All the four countries have a relatively low preference for the data collector or the company that the data is being shared. However, the data collector is the second important attribute in latent class 1. Even though it has relatively low importance for the general public of the Netherlands and Sweden, a group of people valued the data collector highly. Therefore, the data collector and the company that the data is being shared should be considered in sharing digital health information.

The acceptability for data sharing digitally for the best scenario is 89% for Sweden, which is nearly similar to a study conducted on the preference for data linkage for research purpose in Sweden 85.7% (53). The difference could be the acceptability rate calculation method. The Swedish researchers used a conditional logit model, which does not account for individual-level acceptance differences to estimate the acceptability rate. In addition, the difference in the included attributes in the studies could explain the slight difference in acceptability rate. For example, informed consent was not included in their research, which is the most important attribute in this study. On the other hand, the acceptability of data sharing digitally with the best scenario for the Dutch general public is 80.5%, which is lower than the Swedish general public (89%). The observed difference could be explained by the relatively high use of electronic health records by the Swedish general public (74%) compared to the Dutch (44.3%), leading to a positive attitude to share health data digitally.

The marginal change of the attribute level from "not informed" to "being informed" increases the acceptability rate by 25.4 and 28 percentage points for the Dutch and Swedish general public, which is the largest for both countries. The finding is in line with a  study that calculated the marginal change for attribute level change, revealing the biggest marginal change for being informed (46).  The high marginal acceptability rate with informed consent (the most important attribute) is supported by previous research on the general public's attitude on data sharing and preference for data linkage for the research purpose. The researchers noted that the general public needs to be informed when the health information is being shared for research or other purposes (28, 30, 31).

There is a big difference in marginal changes between the Dutch and Swedish general public, especially for "Data shared with" and "Review" attribute level change. For instance, a change of the company that the data is being shared with from a technological company to authority could increase acceptability by 7.7 percentage point for Swedish. Still, for the Dutch, the acceptability is by far low (0.6 percentage point). Generally, the Swedish general public is more responsive for the attribute change than the Dutch general public. Therefore, different acceptability rate would be expected for the Netherlands and Sweden depending on the attribute level change.

## Limitations of the study

When interpreting the findings of this study, the following limitations need to be taken into consideration. Even though DCE is a powerful tool with high positive predictive value to predict real life choices, a hypothetical bias could also affect the estimates. The study used a hypothetical scenario to estimate the population's stated preference, which might not be similar to the actual behaviour on data sharing (54).

The respondents in the study were selected talking account for the age and gender of the populations to have a representative sample. But there is still a significant sociodemographic information difference in the two countries. However, the sociodemographic factors have no significantly determine class membership except for gender which is already stratified.

## Conclusion

The Dutch and Swedish general public do not prefer a technological company as a data collector and do not prefer to share their health information to technological companies for marketing reason and they prefer availability of informed consent and review committee. There is an individual and group level heterogeneity in a preference for sharing digital health data in the Netherlands and Sweden.

# References

1.      WHO. Fifty-eighth World Health Assembly. Revision of the International Health Regulations. 2005.

2.      European Parliament CotEU. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). Official Journal of the European Union. 2016.

3.      Hulsen T. Sharing is caring—data sharing initiatives in healthcare. International journal of environmental research and public health. 2020;17(9):3046.

4.      Haug C. Whose data are they anyway? Can a patient perspective advance the data-sharing debate? New England Journal of Medicine. 2017;376(23):2203-5.

5.      COUNCIL OEJOC. Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine. 1997.

6.      Hermes S, Riasanow T, Clemons EK, Böhm M, Krcmar HJBR. The digital transformation of the healthcare industry: exploring the rise of emerging platform ecosystems and their influence on the role of patients. 2020;13(3):1033-69.

7.      Lovell T. UK government releases details of COVID-19 data-sharing deals with big tech firms after legal action threat. Healthcare IT News. 2020.

8.      WHO. Global Health Observatory (GHO) data, Electronic health records 2016 [Available from: https://www.who.int/gho/goe/electronic_health_records/en/.

9.      WHO. Draft global strategy on digital health 2020–2025. WHO. WHO; 2020 July.

10.     Maiorana A, Steward WT, Koester KA, Pearson C, Shade SB, Chakravarty D, et al. Trust, confidentiality, and the acceptability of sharing HIV-related patient data: lessons learned from a mixed methods study about Health Information Exchanges. Implementation Science. 2012;7(1):1-14.

11.     News B. NHS data breach affects 150,000 patients in England. BBC. 2018.

12.     Manthorpe R. Details of NHS deals with tech giants released by government after legal threat. Sky News. 2020.

13.     Anderson JR, Schonfeld TL. Data-sharing dilemmas: Allowing pharmaceutical company access to research data. IRB: Ethics & Human Research. 2009;31(3):17-9.

14.     Karasneh RA, Al-Azzam SI, Alzoubi KH, Hawamdeh SS, Muflih SM. Patient data sharing and confidentiality practices of researchers in Jordan. Risk management and healthcare policy. 2019;12:255.

15.     Kalra D, Gertz R, Singleton P, Inskip HM. Confidentiality of personal health information used for research. Bmj. 2006;333(7560):196-8.

16.     Ozair FF, Jamshed N, Sharma A, Aggarwal P. Ethical issues in electronic health records: A general overview. Perspectives in clinical research. 2015;6(2):73.

17.     David Blumenthal AC. Apple's Pact with 13 Health Care Systems Might Actually Disrupt the Industry: Healthcare Business & Technology; 2018 [cited 2021 31/01/2021]. Available from: https://www.shrm.org/resourcesandtools/hr-topics/benefits/pages/apple-pact-with-13-health-care-systems-might-actually-disrupt-the-industry.aspx?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+shrm%2Fhrnews+%28SHRM+Online+HR+News%29.

18.     Apple. Use the Health app on your iPhone or iPod touch: Apple; 2021 [Available from: https://support.apple.com/en-us/HT203037.

19.     Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine (1997).

20.     den Exter A. International Health Law and Ethics. Basic Documents2018 Jan.

21.     Essén A, Gerrits R, Kuhlmann EJHp, technology. Patient accessible electronic health records: Connecting policy and provider action in the Netherlands. 2017;6(2):134-41.

22.     Hägglund M, Scandurra I, editors. Patients' Online Access to Electronic Health Records: Current Status and Experiences from the Implementation in Sweden. MedInfo; 2017.

23.     Blok JBaB. Dutch Data Protection Authority sets GDPR fines structure. Engage Hoganlovells. 2019.

24.     GDPR. Data sharing and GDPR compliance: Bounty UK shows what not to do  [Available from: https://gdpr.eu/data-sharing-bounty-fine/.

25.     Norwegian DPA imposes administrative fine to Østfold HF Hospital [press release]. 2020.

26.     Bridges JF, Hauber AB, Marshall D, Lloyd A, Prosser LA, Regier DA, et al. Conjoint analysis applications in health—a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value in health. 2011;14(4):403-13.

27.     Viney R, Lancsar E, Louviere J. Discrete choice experiments to measure consumer preferences for health and healthcare. Expert review of pharmacoeconomics & outcomes research. 2002;2(4):319-26.

28.     Aitken M, Jorre JdS, 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 medical ethics. 2016;17(1):1-24.

29. O'Brien EC, Rodriguez AM, Kum H-C, Schanberg LE, Fitz-Randolph M, O'Brien SM, et al. Patient perspectives on the linkage of health data for research: insights from an online patient community questionnaire. International journal of medical informatics. 2019;127:9-17.

30. Shabani M, Bezuidenhout L, Borry P. Attitudes of research participants and the general public towards genomic data sharing: a systematic literature review. Expert review of molecular diagnostics. 2014;14(8):1053-65.

31. Aitken M, McAteer G, Davidson S, Frostick C, Cunningham-Burley S. Public preferences regarding data linkage for Health Research: a discrete choice experiment. International Journal of Population Data Science. 2018;3(1).

32. Jones RD, Sabolch AN, Aakhus E, Spence RA, Bradbury AR, Jagsi R. Patient Perspectives on the Ethical Implementation of a Rapid Learning System for Oncology Care. Journal of oncology practice. 2017;13(3):e163-e75.

33. Nissenbaum H. Privacy as contextual integrity, 79 Wash. L Rev. 2004;119(121):154-55.

34. Benthall S, Gürses S, Nissenbaum H. Contextual Integrity Through the Lens of Computer Science: Now Publishers; 2017.

35. Ryan M, Bate A, Eastmond CJ, Ludbrook AJBQ, Safety. Use of discrete choice experiments to elicit preferences. 2001;10(suppl 1):i55-i60.

36. Hauber AB, González JM, Groothuis-Oudshoorn CG, Prior T, Marshall DA, Cunningham C, et al. Statistical methods for the analysis of discrete choice experiments: a report of the ISPOR Conjoint Analysis Good Research Practices Task Force. Value in health. 2016;19(4):300-15.

37. Lancaster KJ. A new approach to consumer theory. Journal of political economy. 1966;74(2):132-57.

38. Manski CF. The structure of random utility models. Theory and decision. 1977;8(3):229.

39. de Bekker-Grob E, Donkers B, Bliemer M, Veldwijk J, Swait J. Can healthcare choice be predicted using stated preference data? Social Science & Medicine. 2020;246:112736.

40. Uppsala University UoO, University of Oxford, University of Iceland. Exploring people's knowledge of how electronic health data might be used and preferences for its control [Available from: https://www.cgresearch.org/.

41. contributors W. Netherlands: Wikipedia, The Free Encyclopedia.; 2021 [Available from: https://en.wikipedia.org/w/index.php?title=Netherlands&oldid=1017919210.

42. CBS. Population dashboard [Internet 2021 [Available from: https://www.cbs.nl/en-gb/visualisations/dashboard-population.

43. contributors W. Sweden: Wikipedia, The Free Encyclopedia.; 2021 [Available from: https://en.wikipedia.org/w/index.php?title=Sweden&oldid=1018167113.

44. GmbH S. SurveyEngine GmbH 2021 [Available from: https://surveyengine.com/.

45.		Jennifer Viberg Johansson1* NS, Eik Haraldsdóttir3, Heidi Beate Bentzen4, Sarah Coy2, Jane Kaye2, 5, Deborah Mascalzoni1, 6, Jorien Veldwijk1, 7, 8. Governance mechanisms for sharing of health data: an approach towards selecting attributes for complex Discrete Choice Experiment studies. . Technology in Society, 2021; in press. 2021.

46.		Johansson J, Bentzen HB, Shah N, Haraldsdóttir E, Jónsdóttir GA, Kaye J, et al. Publics' preferences for sharing health data: a discrete choice experiment. 2021.

47.		Johnson FR, Lancsar E, Marshall D, Kilambi V, Mühlbacher A, Regier DA, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR conjoint analysis experimental design good research practices task force. 2013;16(1):3-13.

48.		Jonker M, de Bekker-Grob E, Veldwijk J, Goossens L, Bour S, Rutten-Van Mölken MJJm, et al. COVID-19 Contact Tracing Apps: Predicted Uptake in the Netherlands Based on a Discrete Choice Experiment. 2020;8(10):e20741.

49.		Veldwijk J, Lambooij MS, Kallenberg FG, Van Kranen HJ, Bredenoord AL, Dekker E, et al. Preferences for genetic testing for colorectal cancer within a population-based screening program: a discrete choice experiment. European Journal of Human Genetics. 2016;24(3):361-6.

50.		Pacifico D, Yoo HIJTSJ. lclogit: A Stata command for fitting latent-class conditional logit models via the expectation-maximization algorithm. 2013;13(3):625-39.

51.		Ancillotti M, Eriksson S, Andersson DI, Godskesen T, Fahlquist JN, Veldwijk JJIJoAA. Preferences regarding antibiotic treatment and the role of antibiotic resistance: A discrete choice experiment. 2020;56(6):106198.

52.		Houser KA, Voss WGJRJ, Tech. GDPR: The end of Google and Facebook or a new paradigm in data privacy. 2018;25:1.

53.		Tully MP, Bernsten C, Aitken M, Vass CJBmi, making d. Public preferences regarding data linkage for research: a discrete choice experiment comparing Scotland and Sweden. 2020;20(1):1-13.

54.		de Bekker-Grob EW, Donkers B, Bliemer MCJ, Veldwijk J, Swait JD. Can healthcare choice be predicted using stated preference data? Social science & medicine (1982). 2020;246:112736.

# Appendices

*Additional Table 1: Description of the ranking survey*

| Important factors | Description |
|---|---|
| **Type of information** | Please be aware that "type of information" may include one or more of the following: my health status, my mental health status, genetic information, medication, clinical data, name, DOB, education level, where I live, my ethnicity, political views, religion, sexual orientation, my interests, things I buy, people I see / am friends with,) |
| **Level of identification** | Data can either be identifiable, meaning that your identity can be easily traced or, de-identified, meaning that your identity is protected through some sort of mechanism, a code or other technical means) |
| **Who is the sender or collector of the data?** | The actors that collect data that may then be shared, may be companies, research institutions, healthcare providers etc.) |
| **Recipients** | The actors that receive data after being shared by the sender/collector may be technological or pharmaceutical companies, research institutions, healthcare providers) |
| **Purpose of use of data** | The data may be used for different purposes, such as: commercial advertising or marketing, medical research, policy making, development such as quality improvement or service planning) |
| **Profit-making** | Data may be profitable for different types of entities: companies and/or data subjects) |
| **Who is benefitting when I share the data?** | Sharing the data may result in a benefit for individuals and/or users of data and/or society) |
| **Oversight** | Data sharing may be regulated by oversight bodies carried on by ethic committee, data access committee, national authorities, governmental ministries or departments or groups) |
| | Different types of informed consent: |

| | |
|---|---|
| **Your ability to consent** | - dynamic (you may change your decision over time and you will be informed about changes in the use of data over time<br>- broad (you give your consent once and for all on a very generic purpose: i.e for research)<br>- Specific consent that exactly explains the use and the conditions under which data will be used<br>re-informed consent that will be asked repeatedly for each new use |
| **Soft-law** | (This may include: non-binding guidelines/declarations - usually international declarations, recommendations, and resolutions) |
| **Hard-law** | (This includes: binding normative rules foreseeing: high monetary penalties, other strong penalties of accountability, legally enforced in front of a court) |
| **Public Engagement** | (This may include: public outreach, social activities involving the public in design of policies or representation on boards, surveys about preferences of data sharing, other) |