Privacy Calculus in the Context of the General Data Protection Regulation and Healthcare: A Quantitative Study

Author: Sven Kruthoff
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT,
This research looks at the effect of factors like risk and benefits on the decision of individuals to disclose personal information. As a theoretical framework, the privacy calculus is introduced and connected with the new GDPR regulations. The hypotheses stemming from this connection looks whether there is a significant effect of inherent and handled risk and expected quality of care on the decision to disclose personal information in a hospital setting. This setting revolves around the situation that a participant needs basic or major medical attention. Additionally, a baseline of general attitudes addressing subjects like trust in hospitals or companies and whether personal data is perceived as personal or private is established. The outcome of the survey that included 125 respondents, mainly from Germany or the Netherlands and being between 20 and 30 years old, supports that handled risk and the expected quality of care has a significant influence on the data disclosing decision within subjects. However, since the sample of the data is limited in reach concerning geography and age group no conclusion about the broader population can be made which leaves room for additional research.

Graduation Committee members:
DR. A.B.J.M. WIJNHOVEN,
DR. R.P.A. LOOHUIS

Keywords
Privacy Calculus, GDPR, Hospitals,

11th IBA Bachelor Thesis Conference, July 10th, 2018, Enschede, The Netherlands.

Table of Contents

1. INTRODUCTION

In recent years the further advancement in technology, especially when it comes to gathering data and processing it has rapidly advanced. Data is stored and utilized by companies and data mining tools to utilize users and resources in the best way possible. Especially in consumer services, there is a willing or unwilling collection of data. This does not always lead to further customization and improved service for the data subjects, it often also leaves them vulnerable to the dealings of the data gathering companies. As a consequence, the European Union introduced the new General Data Protection Regulation, that is supposed to protect the subjects whose data is gathered. This regulation does not only affect companies like Facebook but also in hospitals or any other listed company and institutions. The GDPR seeks to improve the rights and the protection of privacy of data subjects, by requiring companies to design around data security and make subjects aware of the data collection and specific use of data. This research will mainly focus on the hospitals and the connected implications of GDPR. In the context of the supposedly improved protection for data subjects, the decisions of how people disclose data might change in the process. The prevalent theory of how the decision to disclose data is described along the lines of privacy calculus. In this research, the primary focus is to make a comprehensive model of the privacy calculus that involves the new EU regulation and other factors relevant to the decision to disclose data. The model of privacy calculus will be adapted in the context of medical procedures in hospitals. The survey is going to sample respondents in the age group between 20 and 30 years old and those who mainly live in Germany or the Netherlands.

2. THEORETICAL FRAMEWORK

2.1 Privacy Calculus in the Medical Field

In the context of data disclosure in the medical field, the theory of privacy calculus is the dominant theory. Privacy calculus describes the decision to disclose data as the tradeoff between benefit and risk of revealing personal data(Knijnenburg et al., 2017). This tradeoff is made to reach the highest utility of another party owning personal information. The utility maximization principle in healthcare seeks for the trade-off between the risk of highly private and confidential data being passed on to maybe unwanted parties and the desire to receive the best possible quality of care. As the disclosed data gains in personal importance, the perceived risk will increase and subjects will choose less to publish data(Malhotra, Kim, & Agarwal, 2004). On the other hand, when people feel that there is little risk involved, they are more likely to disclose personal data. In case people are not aware of the risk of sharing the data, the underlying assumption is not fulfilled. Lacking complete awareness of the risk, people do not expect a reward and hence do not need benefits to persuade them to disclose data. This makes an assumption about a trade-off or an informed decision worthless and disclosing data is not based on the theory of privacy calculus but other random factors.

More specifically in the healthcare sector, the risk is not only concerned about a general breach to the outside in case of an attack on a server and theft of data. It is also about possible errors of the people handling and entering the information or third parties that gain access to data without formal and explicit consent(Dimitropoulos & Rizk, 2009). Because patients do not have the necessary insight into data storage and processing in hospitals, it might lead concerns about the safety of the network(O'Donnell et al., 2011). So a system might be safe even by the highest standards, yet people, due to their lack of knowledge, might not trust it with their data. However, this can also be applied the other way around.

2.1.1 Risk

Users have a general need to protect their personal information(Hui, Teo, & Lee, 2007). Disclosing or sharing data on the internet or in the context of any database exposes them to a certain risk. The five classic dimensions of risk range from being of financial, performance, physical, social and psychological nature(Kaplan, Szybillo, & Jacoby, 1974). In the context of privacy calculus, new dimensions gain importance. Privacy and overall risk are highly relevant when looking at the process of disclosing data because the five classic dimensions do not adequately reflect the new developments and risk factors in online data storage (Featherman & Pavlou, 2003). Additional to the inherent risks of every person, every system consists of the inherent and the handled risk(Bettman, 1973). The inherent risk of a system describes the setup of the system itself and the protection of the data within. The inherent risk of a system is not dependent on the person using it, but on the architecture of the system itself. The handled risk, however, is highly dependent on the user. Different people can assess and handle the risk of data disclosure within a system differently. Meaning that people who have a lot of knowledge about the risk and are adept at handling their data will face less handled risk. Individuals itself are unable to assess risks due to incomplete information and bounded rationality (Krasnova, Kolesnikova, Guenther, & Günther, 2009). In the context of privacy calculus, the real risk of a situation or decision is not included. To make a utility decision, subjects follow their own perceived risk over the actual risks(Khalil & Karam, 2015). To form an opinion about the perceived risk, each individual has, first of all, to recognize the risk itself, assess and estimate it and finally accept the risk (Harbeck, Glendon, & Hine, 2017).

Additionally, to feel the risk of disclosing personal data, people have to value their data. In case they do not perceive it as important or private, they will not associate any risk with it. In case they do not associate any risk with disclosing personal data, individuals do not make a decision based on privacy

calculus. Another factor that influences the risk perception of disclosing the data is the possibility of a breach, exposing personal data to the public or other unwanted recipients. Perceived risk can additionally be divided into institutional trust and concerns for personal privacy (Dinev et al., 2006). Higher trust in institutions and or a lower concern for personal privacy will positively influence the decision to disclose information. Different outside factors influence the perception of risk and benefits and might lead to a different outcome of the choice to disclose information (Kehr, Kowatsch, Wentzel, & Fleisch, 2015). In the context of the data disclosing the trusting beliefs towards the data handling party, can be another reason why the risk perception is different from the actual risk of the decision (Malhotra, Kim, & Agarwal, 2004). The trust factor acts as a mitigating factor when assessing risks. Another mitigating factor influencing risk perception is the perceived fairness of the data collection. Those factors include data collection within an existing context, the ability to control the data, the relevance of the data and data will be used to draw correct and reliable conclusions about your situation (Culnan & Armstrong, 1999).

2.1.2 Benefits

The benefits, countering the risks of disclosing personal data, have to, as said before, have to be bigger to lead to a decision to give up personal data. Factual benefits of publishing personal data in the information technology healthcare context are improved and more efficient care and an improved transferability between doctors (Glaser, Henley, Downing, Brinner, & Community, 2008). Additionally, a more digitized, information systems oriented healthcare will reduce errors and improve efficiency (Noffsinger & Chin, 2000). Again, it is important to stress that the benefits might only be perceived or anticipated as such and subjectively estimated by individuals. Perceived benefits influence the knowledge about the handling of the data and the connected service and the personal desire for personalization of the service (Gómez-Barroso, Feijóo, & Martínez-Martínez, 2018). A lack of knowledge concerning the handling of the data and the process itself will lead to an overestimation of the benefits. The same counts for the desire to have a customized service.

2.2 General Data Protection Regulation

In the context of this research the new EU directive: General Data Protection Regulation (GDPR) gains relevance since it directly influences specific factors in the privacy calculus process. The GDPR overhauls the long obsolete Data Protection Directive 95/46/EC, which was established in 2005. The GDPR is mandatory not only for all companies that are based in the EU but also those offering services in the EU, meaning that also multinational companies like Facebook have to comply whit its regulations. The six focal points of GDPR are the breach notification, right to access, right to be forgotten, data portability, privacy by design and the role of a data protection officer. Establishing a data protection officer is less relevant for this research as it has less impact on the user itself. The breach notification constitutes the need for a company to immediately disclose to clients, in case of a breach, where personal data could have been compromised. The notification has to be sent within the first 72 hours of being made aware of the breach. The right to access states that data subject can request information about their data and the way it is being processed at any point. The data handler has to provide that data in an electronic format and free of charge. The right to be forgotten means that data subjects can request to have their personal data deleted from company records and not be processed as a subject in any other way. Data portability ensures that all data that was collected can be put out in a machine-readable format and transferred to other systems. In the context of the new GDPR, privacy by design means that the protection of individual data has to be at the center when setting up information systems. Before that, it was mostly only seen as a necessary add-on to these. The new laws on data protection will make the theory of privacy calculus more critical, as people are made aware of what is done with their data and how it is processed. Hence this will give rise to a more conscious decision of disclosing data to a company or institution, due to an increase in awareness that risk is taken when disclosing data.

2.3 Hypothesis

Based on the theory of privacy we come up with the following model:

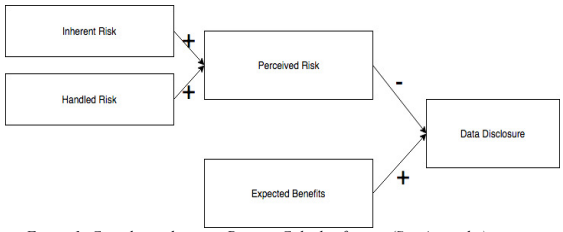

Figure 1: Correlation between Privacy Calculus factors (See Appendix)

The underlying assumption of the depicted model is that people are aware of the risk of disclosing personal data to third parties. If they are not aware of the risk, the decision is not based on privacy calculus, but other random factors, making this research obsolete. So this model depicts a positive relationship between inherent and handled risk towards the perceived risk of an individual. So increasing the handled and inherent risk increases the perceived risk of the subject. In case the security of a system increases, inherent risk decreases, so the perceived risk decreases too. In case that an individual has more knowledge about privacy and is better at handling the risk, the handled risk decreases and consequently the perceived risk decreases. The perceived risk is negatively correlated with the decision to disclose

data, meaning that a high risk will lead to little willingness to give up personal data. Expected benefits are positively correlated with the decision to disclose data, as it makes people more willing to take risky decisions. Based on these correlations we establish the following hypotheses:

H1: People perceive their personal data is important.
This hypothesis stems from the basic requirement of people to perceive their data as important or valuable to make an informed privacy calculus decision. In case they do not perceive their data as important, they would not associate any risk with the disclosure of personal data, making the application of the privacy calculus theory obsolete. This is used as a baseline to recognize possible outliers and see if the theory of privacy calculus applies to the number of respondents. In case this condition is not met, the answers in the survey are not based on privacy calculus.

H2: People place higher trust in hospitals than they place in commercial companies.

Institutional trust plays a prominent role as a mitigating factor for risk perception, meaning that a hospital is an institution people associate more trust with than companies. In the context of privacy calculus, people will then associate less risk with disclosing information to a hospital, meaning they are more likely to give personal data over to hospitals. The more significant trust is also measured in the baseline by comparing the trust in companies and hospitals.

H3: People facing a more severe condition, are more likely to share personal data than people facing a minor condition
.
Based on the theory of privacy calculus, people facing a more severe condition expect or need a higher level of quality of care, meaning that they are more willing to take the risk. This risk is presented in the form of giving up private information, not only to the treating physician but also others that can help. To test the hypothesis, comparing the results of t-tests between heavy and light diseases in each different scenario is used.

H4: When introducing the GDPR, people are more likely to disclose personal data.

Introducing the GDPR decreases the level of inherent risk in the system itself, by introducing privacy by design. Decreasing the inherent risk would lead to a less risky perception of another party having the data, likely resulting in less concern when disclosing data. Introducing GDPR, will only reduce the risk and not the benefits. As people try to maximize personal utility of disclosing data, reducing the risk while keeping the same level of benefits increases utility to the disclosing party. To say whether we support this hypothesis, pairing the light diseases in the basic scenario and the GDPR scenario and pairing the heavy diseases in the same way. In case we find a significant difference in the samples we acknowledge the hypothesis as true.

H5: When giving the choice to opt-out of a decision, people are more likely to disclose personal data.

Opting-out represents a decrease in handled risk in the context of the privacy calculus decision. When giving the chance to withdraw from data storing, people will feel that the decision made is less risky and should subsequently lead to higher disclosing behavior. As done in the hypothesis before, the method is to pair light diseases in each scenario and separately and then look at the results, if we can find a significant difference in the two distributions.

H6: When suggested that sharing more data with the hospital, will lead to a higher level of care, people are more likely to disclose data.

Suggesting that sharing more data with the hospital will result in better care increases the benefits of taking that risk. In order of that increase, people should be more inclined to disclose data. Again, we test this hypothesis by separately comparing light and heavy diseases in every scenario.

3. METHODOLOGY

3.1 Basic Methodology

The survey is designed to measure the causal relationships of the privacy calculus applied to the medical field. The study answered by peers and data is collected via an online survey. Utilizing an online survey is the best choice at hand because it is cheap, quick and it is easily accessible for everyone. Additionally, in the context of their personal space, in which the survey is most likely answered, the answers to the questions will be given truthfully, as the data subjects do not face any peer or social pressure. The online survey is measuring the perceived risk and the expected quality of the care. The perceived risk is divided into the inherent risk and the handled risk of the system. On the other side, the expected quality of care is measured. First, a baseline of the intention to disclose data is taken, meaning that under no outside conditions, what kind of data would be disclosed by the respondent. The dependent variable is the actual decision to disclose an amount of data and is measured in 4 different levels. Those levels vary from not disclosing any data, over disclosing medical data only to the treating medical staff. The next level would be to give the medical staff full authority to act in the patients' best interest, possibly passing on the information to other medical experts. The fourth level is not only giving the medical staff full authority over personal data, but also provide consent to store the data so it can be used in longitudinal studies, to improve treatment of following patients. The four levels show a different level of data disclosure on an ordinal scale and act as

a slider, where with every additional step more access to data is granted.
At first, a baseline of perceived risk towards the handling of data by companies and the personal valuation of data is measured. The baseline is measured on a 5 point Likert scale ranging from "strongly disagree" to "strongly agree." In the further analysis, a score of 1 corresponds with the answer of "strongly disagree," and a score of 5 corresponds with the answer of "strongly agree." The six concepts we measure is the perception, whether personal data is private, or essential to the data subject. Then we ask about the trust in companies in general. Afterward, we ask for the trust in medical institutions specifically and whether there is trust in doctors not to take advantage of personal data. Then it is asked if monetary incentives could persuade them to share their data. The 7th question is an inverted variable, meaning that it is a question to see whether the baseline questions are answered consistently. This baseline is essential to measure, to explain the possible variance between answers of subjects. Additionally, it is important in the theory of privacy calculus to realize the level of risk involved to make a conscious decision about disclosing data. In case people do not see the risk in sharing the data, as before mentioned they do not make a privacy calculus decision, and therefore are not relevant to this research. To make relevant assumptions about the correlation and relations between we need more than 100 respondents (Cooper & Schindler, 2014). Otherwise, tests like the t-test give us no significant answer about the difference between two variables, because the means and variations would be too imprecise.
After that, we expose the participants to possible situations they could come across. These situations include changing only one of either perceived risk or expected quality of care because, in the case of altering more than one variable, it is not possible to attribute the change in the dependent variable to one of the independent variables exactly. In the first situation, participants are presented with a minor health issue, requiring basic surgery. In this situation, it is measured by what kind of data they are willing to disclose. The second situation involves a major disease, for example, cancer and again the level of disclosing information is measured. The basic conditions are changed to see if there is a difference in disclosing data when it comes to the different severity of diseases.
The risk is divided into the inherent risk and the handled risk. The participants are going to be presented with the situation to opt in and out of disclosing the data, representing the ability to handle the risk. In the following question, participants are made aware of the new GDPR regulation and the following improvements to data security within the system. Then again the type of data they are willing to disclose is measured. The improvement in the data security through GDPR represents the inherent risk of the system. Both factors are expected to increase the willingness to disclose information about themselves. The third situation respondents are presented with is the quality of care that is expected to be given by the hospital, meaning that disclosing data means better care. Again, it is expected to disclose more data in case of a more severe disease.

3.2 Survey questions

Perceived Risk
On a 5 point Likert scale would you say that:
- your personal data is important
- your personal data is private
- you trust companies with your data
- hospitals are a trustworthy institution to store data
- you trust doctors to not take advantage of your data
- monetary incentives could persuade you to share more data?
- you trust hospitals more with your personal data than commercial firms?

Light Disease
You are admitted into a hospital with a broken arm. What personal data are you giving to the physician?

Bad Disease
You are admitted into a hospital with a yet unknown disease, showing cancer symptoms. What personal data are you giving to the physician?

Inherent risk
Under the new EU legislature for the General Data Protection Regulation, hospitals are required to improve the security and the procedure design of internal data handling.

You are admitted into a hospital with a broken arm. What personal data are you giving to the physician? You are admitted into a hospital with a yet unknown disease, showing cancer symptoms. What personal data are you giving to the physician?

Handled Risk
After giving the data to the hospital, you have the option to opt out. This means that at any time after the treatment you can choose to have your data erased from any database.

You are admitted into a hospital with a broken arm. What personal data are you giving to the physician? You are admitted into a hospital with a yet unknown disease, showing cancer symptoms. What personal data are you giving to the physician?

Expected Quality of Care
The hospital reports that they cannot guarantee appropriate quality of healthcare if they do not possess the necessary information from your medical history.

You are admitted into a hospital with a broken arm. What personal data are you giving to the physician? You are admitted into a hospital with a yet unknown disease, showing cancer symptoms. What personal data are you giving to the physician?

Since all the questions are measured on an ordinal scale, the data is going to analyze by using a paired t-test. With the t-test, it is possible to see, if the answers in the pair differ from each other significantly. The t-test does this by comparing the null hypothesis that there is no relationship between the two samples. The level of significance chosen in this research is 5% two-tailed to reject the null hypothesis, meaning that there is a relationship between the samples and this difference is accounted for by the change in the variables at hand. The t-test is based on the assumption of normality, meaning that the data has to be normally distributed for the test to be viable. To measure the normality, we look at the skewness of each variable. The cut-off values for skewness are between 1 and -1 because everything outside of the values is considered strong skewness, making it impossible to assume normality.

4. DATA ANALYSIS

4.1 Description of Results

At the end of the data collection process, 125 answers to the survey were collected. The majority of the respondents come from either the Netherlands or Germany, while the rest except 2 came out of the European Union. The distribution between male and female is 56% to 44%. 25 out of the 125 answers were not between the age of 20 and 30. This puts the sample right in the scope of the research which is aimed at peers in the same age group and living in the European Union, where the new GDPR takes effect.

Statistics

		Pers_Data_Imp	Pers_Data_Priv	Trust_Companies	Trust_Hospitals	Trust_Doctors	Value_Custom	Incentive_Sharing	Med_Hist_Research	Hosp_Over_Comp
N	Valid	125	125	125	125	125	125	125	125	125
	Missing	0	0	0	0	0	0	0	0	0
Mean		4.11	4.00	2.75	3.58	3.97	3.92	2.64	3.43	4.02
Median		4.00	4.00	3.00	4.00	4.00	4.00	3.00	4.00	4.00
Std. Deviation		.935	1.078	1.029	1.002	.975	1.060	1.073	1.234	.954
Minimum		1	1	1	1	1	1	1	1	1
Maximum		5	5	5	5	5	5	5	5	5

Figure 2: Means of Baseline measure (See Appendix)

The results of the first part of the survey, the baseline as depicted above, yields the results, that respondents agree that their data is personal and important (mean of 4.11 and 4.00). The respondents expressed having higher trust in hospitals, when it comes to handling their data, while they express lower trust in companies on the same subject. They even express higher trust in doctors to not misuse their data compared to hospitals. Even if the subjects value the possible customization of disclosing data, they are not willing to act on the opportunity to share their data when receiving monetary incentives for it. The last variable supports what is displayed in the 3rd and fourth variable, meaning that respondents place higher trust in hospitals than they place in companies. This supports the claim that people answered the questions consistently and not at random.

To see whether we can apply the t-test, we check the normality of the variables by looking at the skewness of the data. We find that none of the values are outside the chosen interval of 1 and -1 (See table 1 in Appendix). So we can assume that the data is normally distributed and we can apply the t-test to test the significance of the difference in variables.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Basic_Light - Basic_Heavy	-.760	.712	.064	-.886	-.634	-11.940	124	.000
Pair 2	GDPR_Light - GDPR_Heavy	-.704	.696	.062	-.827	-.581	-11.311	124	.000
Pair 3	Opt_Light - Opt_Heavy	-.664	.761	.068	-.799	-.529	-9.750	124	.000
Pair 4	Quality_Light - Quality_Heavy	-.560	.689	.062	-.682	-.438	-9.092	124	.000

Figure 3: Paired Sample T-Test for Basic-Heavy conditions (See Appendix)

The results of the paired t-test in the different scenarios are applied to see if the response to the different situations differs based on the type of disease. The t-test tests for the difference in means grouped around the variance of the distributions and depicts if the hypothesis that both distributions are not the same is correct. When comparing the two sterile scenarios of the light and the heavy disease, it is seen that the difference between the light and the heavy disease is negative and also significant on the given 5% level. A negative difference is meaning that respondents tend to disclose more data when facing a heavy disease. The same phenomenon is found in the other three situations. So all tests show that we reject the null hypothesis that the answer distribution is not the same.

The same measures as a paired t-test and correlation analysis are used to compare the different scenarios and the different diseases in the scenarios.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Basic_Light - GDPR_Light	-.072	.443	.040	-.150	.006	-1.816	124	.072
Pair 2	Basic_Heavy - GDPR_Heavy	-.016	.458	.041	-.097	.065	-.391	124	.697
Pair 3	Basic_Light - Opt_Light	-.232	.649	.058	-.347	-.117	-3.995	124	.000
Pair 4	Basic_Heavy - Opt_Heavy	-.136	.600	.054	-.242	-.030	-2.533	124	.013
Pair 5	Basic_Light - Quality_Light	-.272	.734	.066	-.402	-.142	-4.145	124	.000
Pair 6	Basic_Heavy - Quality_Heavy	-.072	.709	.063	-.198	.054	-1.135	124	.258

Figure 4: Paired Sample T-Test for Basic-Scenario Conditions (See Appendix)

When looking at the paired test for the effect on the scenarios, the only significant pair of compared distributions is found between the sterile conditions and the option to withdraw from the data disclosing decision. Only when looking at the light disease in the case of the basic conditions and the suggestion that all data is needed to give the best treatment, we find a significant difference in disclosing behavior. The rest of the tests fail to reject the null hypothesis on the chosen significance level of 5%.

The correlation between the light and heavy diseases in the same scenarios supports the findings of the t-tests. We only find moderate correlations when looking at Spearman's Rho, supporting that the distributions are not the same. If the variables scored high on correlation, we would find low significance on t-test measures. This is seen in the correlation between the basic scenario and GDPR scenario when looking at the heavy disease. While the correlation is high (0.801) we find the lowest significance out of all the pairs (0.697).

4.2 Hypothesis analysis

H1: People perceive their personal data is important

As seen in the baseline measures, the respondents agree with the statement that their data is private and important (Mean *Pers_Data_Imp:* 4.11; Mean *Pers_Data_Priv*: 4.00). So we can conclude that the following answers can be based on the underlying assumptions made to decide, whether to disclose data or not, is based on the privacy calculus framework.

H2: People place higher trust in hospitals than they place in commercial companies.

The answers to the questions, whether there is a difference in trust between commercial companies and hospital gives the conclusive report to support the hypothesis. The mean for trust in hospitals is higher than the trust in companies (Mean *Trust_Hospitals:* 3.58; Mean *Trust_Companies:* 2.75). Additionally, a t-test rejects the null hypothesis, that both samples are the same, at a significance level of 5%. The control variable of trust in hospitals over trust in companies supports the hypothesis too (Mean *Hosp_Over_Comp:* 4.02).

H3: People facing a more severe condition, are more likely to share personal data than people facing a minor condition.

Based on the paired t-test between the heavy and light disease in each of the four scenarios, we always find a significant negative difference, meaning that people facing a more severe condition share more information than those who face a minor condition. So we can say, that with these responses, an increased level of expected quality of care influences the decision to disclose data in the way we expected.

H4: When introducing the GDPR, people are more likely to disclose personal data.

Based on the testing done in the analysis, we fail to confirm the hypothesis. First of all, the difference between both of the pairs is not significant on a 5% level, but also within the 95% confidence interval, we can not even say if the direction is positive or negative. So based on the research we can conclude that GDPR and subsequently the inherent risk do not have a significant impact on the disclosing behavior of people.

H5: When giving the choice to opt-out of a decision, people are more likely to disclose personal data.

The t-test when comparing the basic scenario and given a choice to opt out gives us the conclusion that the choice to opt has a significant influence on the decision to disclose information. This means that the handled risk has a significant influence in the context of privacy calculus and disclosing data.

H6: When suggested that sharing more data with the hospital, will lead to a higher level of care, people are more likely to disclose data.

Testing this hypothesis, the paired t-test shows that only the difference in distributions in the light disease is significant on a 5% level. The comparison in the heavy disease does not remotely show a significant difference between both answer distributions (p=.258). This shows that only on a level where there would be less risk-taking behavior, the suggestion of a hospital to disclose as much data has a significant influence. So conflicts with the second hypothesis, because they both address the expected level of care.

5. DISCUSSION

After conducting the tests, it becomes clear, that not all the hypotheses turned out to be true on a significant level. There is no evidence that in case of a heavy disease, the suggestion of hospitals to give more data for better treatment has a significant change. This might be because of an already high willingness to accept the risk when it comes to a possibly terminal disease. So people might already be at their personal limit for disclosing data, and additional expected benefits might not provide a higher utility in case of data disclosure. However, the change when looking at a light disease stems from a high trust in hospitals as shown in the baseline. So in the eyes of a respondent suggestions by a trustworthy source, like the hospital, are more likely be believable and as a consequence subjects act upon the suggestion.

Against expectations, the introduction of the GDPR did not change the disclosing behavior, making clear, that a change in the inherent risk does not have a significant impact on the risk perception in this sample. Another way of explaining it could be that people do not realize the impact the introduction of the new GDPR has on the actual safety of their data, as they are not experts on the subject. However, this is a less valid point, because privacy calculus argues about the perceived risk and not the actuals risk are taken into account. It is possible that due to the press coverage around the introduction of the GDPR, the new regulation is more present, but people are doubtful, whether to believe its' usefulness.

5.1 Limitations of the research

The research itself presents a valid insight into the opinion and effects of privacy calculus on people between 20 and 30 in middle Europe especially, Germany and Netherlands. The dataset features more than 100 respondents in the relevant categories. Based on the literature a sample of 100 people is sufficient enough to make correct assumptions about the population (Cooper & Schindler, 2014). The answers might significantly differ in other regions of Europe that are less developed when it comes to technology and especially other age groups. People older than the chosen group are very likely to have a different perception of technology, and its safety. Future research can pick up this topic and expand the

same model onto an older age group, to find if there is a significant difference between the two groups.

5.2 Contributions to practice

This research presents a good picture of the state of privacy calculus within the given age group and geographic context, especially since the new GDPR directive is just introduced. This could be a benchmark for future research to compare the perceived risk for future research. Primarily, the baseline can be used in the future, because the further changes made to law revolving around data security and data subject protection might change the views on companies and the perceived privacy of data. Additionally, the research done offers a simple approach to the theory of privacy calculus. Additionally, this research can be taken up as a starting point for further research on privacy calculus in the medical field. What we have seen in the survey, is that possible terminal diseases seem to override the basic assumptions that benefits have to be presented in order to disclose data.

BIBLIOGRAPHY

Glaser, J., Henley, D. E., Downing, G., Brinner, K. M., & Community, F. the P. H. C. W. of the A. H. I. (2008). Advancing Personalized Health Care through Health Information Technology: An Update from the American Health Information Community's Personalized Health Care Workgroup. *Journal of the American Medical Informatics Association*, *15*(4), 391–396. Retrieved from http://dx.doi.org/10.1197/jamia.M2718

Noffsinger, R., & Chin, S. (2000). Improving the delivery of care and reducing healthcare costs with the digitization of information. *Journal of Healthcare Information Management : JHIM*, *14*(2), 23–30.

Cooper, D. R., & Schindler, P. S. (2014). *Business Research Methods* (12th ed.). Boston: McGraw Hill/Irwin.

Noffsinger, R., & Chin, S. (2000). Improving the delivery of care and reducing healthcare costs with the digitization of information. *Journal of Healthcare Information Management : JHIM*.

Hui, K.-L., Teo, H. H., & Lee, S.-Y. T. (2007). The Value of Privacy Assurance: An Exploratory Field Experiment. *MIS Quarterly*, *31*(1), 19–33. https://doi.org/10.2307/25148779

Bettman, J. (1973). Perceived Risk and Its Components: A Model and Empirical Test. *Journal of Marketing Research*, *10*, 184–190.

Kaplan, L. B., Szybillo, G. J., & Jacoby, J. (1974). Components of perceived risk in product purchase: A cross-validation. *Journal of Applied Psychology*, *59*(3), 287–291. https://doi.org/10.1037/h0036657

Krasnova, H., Kolesnikova, E., Guenther, O., & Günther, O. (2009). "It Will not Happen To Me!": Self-Disclosure in Online Social Networks. *Amcis 2009 Proceedings*, 343. https://doi.org/10.7892/boris.47460

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet Users ' Information Privacy Concerns (IUIPC): The Construct , the Scal ... https://doi.org/10.1287/isre.1040.0032

Featherman, M. S., & Pavlou, P. A. (2003). Predicting e-services adoption: A perceived risk facets perspective. *International Journal of Human Computer Studies*, *59*(4), 451–474. https://doi.org/10.1016/S1071-5819(03)00111-3

Khalil, L., & Karam, N. A. (2015). Security Management: Real versus Perceived Risk of Commercial Exploitation of Social Media Personal Data. *Procedia Computer Science*, *65*(Iccmit), 304–313. https://doi.org/10.1016/j.procs.2015.09.087

Culnan, M. J., & Armstrong, P. K. (1999). Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. *Organization Science*, *10*(1), 104–115. https://doi.org/10.1287/orsc.10.1.104

Dinev, T., Bellotto, M., Hart, P., Russo, V., Serra, I., & Colautti, C. (2006). Privacy calculus model in e-commerce - A study of Italy and the United States. *European Journal of Information Systems*, *15*(4), 389–402. https://doi.org/10.1057/palgrave.ejis.3000590

Gómez-Barroso, J.-L., Feijóo, C., & Martínez-Martínez, I. J. (2018). Privacy calculus: Factors that influence the perception of benefit/ Cesión calculada de información personal: factores que influyen en la percepción de beneficio. *El Profesional de La Información*, *27*(2), 341–348. https://doi.org/10.3145/epi.2018.mar.12

Kehr, F., Kowatsch, T., Wentzel, D., & Fleisch, E. (2015). Blissfully ignorant: The effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal*, *25*(6), 607–635. https://doi.org/10.1111/isj.12062

O'Donnell, H. C., Patel, V., Kern, L. M., Barrón, Y., Teixeira, P., Dhopeshwarkar, R., & Kaushal, R. (2011). Healthcare consumers' attitudes towards physician and personal use of health information exchange. *Journal of General Internal Medicine*, *26*(9), 1019–

1026. https://doi.org/10.1007/s11606-011-1733-6

Dimitropoulos, L., & Rizk, S. (2009). A state-based approach to privacy and security for interoperable health information exchange. *Health Affairs*, *28*(2), 428–434. https://doi.org/10.1377/hlthaff.28.2.428

Davis, F. D. (1986). A technology acceptance model for empirically testing new end-user information systems: Theory and results. *Management*, *Ph.D.*(April), 291. https://doi.org/oclc/56932490

European Parliament. (2016). General Data Protection Regulation 2016/679. *EUR Lex*, *0011*(April), 1–341. https://doi.org/http://eur-lex.europa.eu/pri/en/oj/dat/2003/l_285/l_285 20031101en00330037.pdf

Knijnenburg, B. P., Raybourn, E., Cherry, D., Wilkinson, D., Sivakumar, S., & Sloan, H. (2017). Death To The Privacy Calculus? Retrieved from https://poseidon01.ssrn.com/delivery.php?ID=421000089122085077069092121095096031057007068045058024089023086126083022086087028010033060122005052003049115074023007082104085112008078077085086093020108069074107124000078053121079084125119123005002089114065065 0

Harbeck, E. L., Glendon, A. I., & Hine, T. J. (2017). Reward versus punishment: Reinforcement sensitivity theory, young novice drivers' perceived risk, and risky driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *47*, 13–22. https://doi.org/10.1016/j.trf.2017.04.001

APPENDIX

TABLE 1

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Basic_Light	125	1	4	2.45	.798	.173	.217
Basic_Heavy	125	1	4	3.21	.687	-.598	.217
GDPR_Light	125	1	4	2.52	.768	.203	.217
GDPR_Heavy	125	1	4	3.22	.633	-.414	.217
Opt_Light	125	1	4	2.68	.829	-.030	.217
Opt_Heavy	125	1	4	3.34	.649	-.659	.217
Quality_Light	125	1	4	2.72	.867	-.025	.217
Qualtiy_Heavy	125	1	4	3.28	.758	-.971	.217
Valid N (listwise)	125						

FIGURE 1

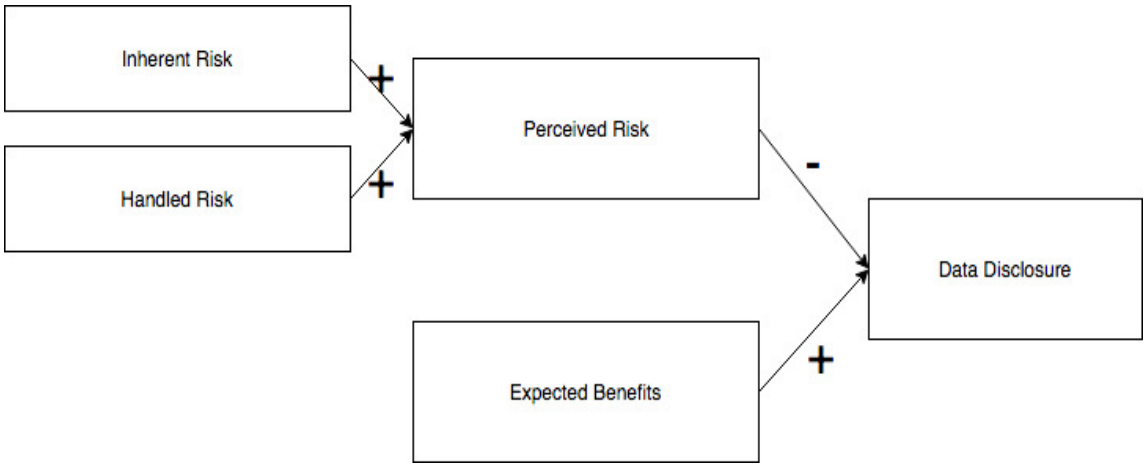

FIGURE 2

Statistics

		Pers_Data_Imp	Pers_Data_Priv	Trust_Companies	Trust_Hospitals	Trust_Doctors	Value_Custom	Incentive_Sharing	Med_Hist_Research	Hosp_Over_Comp
N	Valid	125	125	125	125	125	125	125	125	125
	Missing	0	0	0	0	0	0	0	0	0
Mean		4.11	4.00	2.75	3.58	3.97	3.92	2.64	3.43	4.02
Median		4.00	4.00	3.00	4.00	4.00	4.00	3.00	4.00	4.00
Std. Deviation		.935	1.078	1.029	1.002	.975	1.060	1.073	1.234	.954
Minimum		1	1	1	1	1	1	1	1	1
Maximum		5	5	5	5	5	5	5	5	5

FIGURE 3

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Basic_Light - Basic_Heavy	-.760	.712	.064	-.886	-.634	-11.940	124	.000
Pair 2	GDPR_Light - GDPR_Heavy	-.704	.696	.062	-.827	-.581	-11.311	124	.000
Pair 3	Opt_Light - Opt_Heavy	-.664	.761	.068	-.799	-.529	-9.750	124	.000
Pair 4	Quality_Light - Qualtiy_Heavy	-.560	.689	.062	-.682	-.438	-9.092	124	.000

FIGURE 4

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Basic_Light - GDPR_Light	-.072	.443	.040	-.150	.006	-1.816	124	.072
Pair 2	Basic_Heavy - GDPR_Heavy	-.016	.458	.041	-.097	.065	-.391	124	.697
Pair 3	Basic_Light - Opt_Light	-.232	.649	.058	-.347	-.117	-3.995	124	.000
Pair 4	Basic_Heavy - Opt_Heavy	-.136	.600	.054	-.242	-.030	-2.533	124	.013
Pair 5	Basic_Light - Quality_Light	-.272	.734	.066	-.402	-.142	-4.145	124	.000
Pair 6	Basic_Heavy - Qualtiy_Heavy	-.072	.709	.063	-.198	.054	-1.135	124	.258