Effectiveness of phishing in healthcare: A meta-analysis

Philip Khasuntsev
p.khasuntsev@student.utwente.nl
BSc Health Sciences
Enschede, Netherlands

Abstract

Phishing attacks are increasingly becoming a bigger problem every year. In the United States, an average data breach will cost 3.86 billion dollars. Especially phishing emails are common to use among criminals. In healthcare, a lot of sensitive data is stored in the EHR, which is interesting for criminals. Hospitals that are victimized by phishing emails, could potentially no longer treat patients because there is no access to the EHR. In this paper, a meta-analysis is performed to understand what influence the characteristics of an individual, the characteristics of the phishing email itself, and the context alignment have on the susceptibility of a phishing email. The results do not proof a significant effect from the characteristics that are used for this meta analysis.

Keywords: Email, effect, meta-analysis, persuasion principles, phishing

Samenvatting Phishing aanvallen worden een steeds groter probleem. In de Verenigde Staten kost een gegevenslek gemiddeld 3.86 miljard dollars. De phishing email is een vaak gebruikte methode voor criminelen. In de gezondheidszorg wordt veel informatie opgeslagen in het EPD, wat dit interessant maakt voor internet criminelen. Zorginstellingen die slachtoffer worden van een phishing aanval en daarbij toegang tot het EPD kwijtraken kunnen mogelijk geen zorg meer verlenen aan patiënten. In dit onderzoek is een meta-analyse uitgevoerd, waarbij is gekeken naar het effect op van verschillende factoren van een individu of van de phishing mail zelf, op de effectiviteit van de email. Ook de context waarin de mail is verstuurd, is onderzocht. Uit de meta-analyse is gebleken dat er geen significante effecten zijn van de factoren die zijn onderzocht.

1 Introduction

90% of all the organizations worldwide have been targeted by cybercriminals and have received phishing emails. Recent studies show us that this type of cybercrime is becoming an ever-increasing problem for modern-day society, enabling criminals to gain access into information systems [3]. The financial damage caused by these data breaches as a consequence of phishing emails has dramatically increased in the past years, where in 2021 the average cost of data breaches in the United States has reached 3.86 billion dollars[28].

An email created to obtain credentials or steal the user's information is called a phishing email[4]. In the literature, different definitions have been used for phishing emails. A systematic literature review has combined this into one definition: "Phishing is a scalable act of deception whereby impersonation is used to obtain information from a target."[21]. This email could contain a link that will redirect the user to a duplicated login page that the user is used to seeing, but this malicious website will store the username and password and sell, or even use them to get access into the information system. When attackers are designing the email for a specific goal and population of a company, this is called a spear attack[3].

Social engineering is a non-technical method where criminals attempt to attack humans in order to break into an information system. Social engineering and phishing emails are often used together, where attackers manipulate the receiver and try to let them take action to give access to the attacker into the system[19]. The ultimate goal of using social engineering in combination with a phishing attack is to manipulate the victim. Social engineering is a step above phishing, it is a method that could be used in combination with other methods as well, called phishing for example. An example of using social engineering in combination with a phishing email is using the real name of a victim or closing the email with the name of their manager. This way criminals are trying to let the receiver believe the email is legitimate. Once the receiver falls for the phishing email, the attacker could have gained access to data, and the breach is completed.

In healthcare organizations, information about patients is stored in the Electronic Health Record (EHR). This data contains a lot of sensitive information, such as name, address, and contact information. In addition, it can also contain information about the treatment plan or laboratory results. This EHR should only be accessible to authorized healthcare employees and should always be available and correct. Criminals are interested in the EHR for two main reasons, to sell personal information about patients or to encrypt the EHR and ask for large amounts of money to decrypt it again[20]. The latter has a direct impact on patient safety, since without the important information of the EHR, patients cannot be

treated. Besides the importance of the EHR, the EHR should also be accessible to people with multiple functions: doctors, nurses, secretaries, or administrative employees. Each of these users is a potential target for phishing attacks, resulting in a larger risk of breaching the information system.

However, not every person has the same likelihood to get victimized by a phishing attack, the characteristics of a person determine the risk[23]. Sex, age, and awareness of phishing attacks are examples of such characteristics. Knowing the specific characteristics of a person results in a different risk of becoming a victim of such attacks, phishing awareness trainings could be focused more on those people, resulting in an information system that is more resistant to phishing.

The target is not the only factor that influences the risk of falling for an attack, the phishing mail itself has a major impact on the risk. Attackers can use many different techniques to manipulate the target, for example sending a fake email impersonating the government [23]. What are these characteristics, and which characteristics are more likely to make the phishing attack successful? Sharing these techniques during a training will help people to recognize a phishing email more effectively.

Finally, attackers carefully think about the context of the phishing email[12]. This is where social engineering is used. Because the content of the email should be in alignment with the target's situation, getting an email from a country you have never been to would most likely be ignored by the target. If we know which contexts are used mostly and are effective, the chances of a successful phishing attack can be reduced.

The general research question of this thesis is:

How do individual characteristics of people, attributes of a phishing email, and contextual factors influence the effectiveness of phishing email attacks within healthcare organizations?

To formulate an answer, the following sub-research questions will be answered in this thesis:

1. What is the influence of individual characteristics of healthcare employees on the effectiveness of phishing email attacks?
2. What is the influence of characteristics of a phishing email on the effectiveness of this attack?
3. What is the influence of context alignment of a phishing email on the effectiveness of this attack?

2 Theoretical framework

2.1 Social engineering

Social engineering is the art of influencing and manipulating people with the intention of gaining access to specific information systems or just intending to steal sensitive information. With social engineering, attackers try to let the victim perform actions they usually won't do, for example downloading an attachment or clicking on a link and entering their username and password [28]. Social engineering is not only limited to the digital world but it can also be used in person, for example, to obtain physical keys from healthcare employees. The latter is an example of direct social engineering[31]. When using social engineering in combination with phishing email, this is a non-real-time type of social engineering [31]. As we can see, social engineering is an element of phishing attacks. By using social engineering, the attackers are designing the mail in such a way, that the effectiveness will increase. Attackers usually also try to appeal to the emotions of the victim, for example by inducing fear or creating excitement. This will more likely cause a successful phishing attack, which can result in breaching into an information system.

2.2 Phishing email

One of the most used cyber-attacks is the phishing email[13]. Such an email is designed to steal credentials or obtain sensitive information about the victim by using a fake email address or redirecting to a fake website. Attackers are interested in passwords, usernames, or specific information. Another goal is to infect the system with malware. Previous research has shown that 30% of phishing emails have been read by the victim, another 12% of those victims even clicked on the link in the email or performed actions, for example opening malicious attachments [11]. Phishing emails are focused on individuals because those people are the weak point in the information system. Most companies are providing trainings for their employees, such a training which will help employees to recognize a phishing email, and how to properly react to such a mail (reporting it for example)[22]. There are various types of phishing emails. Attackers can send a phishing email to a lot of people at the same time, without specific context for an individual. Sometimes this can already be a sign of a phishing attack, for example when an employee of hospital X receives a fake email from hospital Y. When using social engineering, it is possible to send a more personal phishing email to employees of a company, using the right name of a manager or about a current project that the employee is working on. Such a more specific attack is also known as a spear phishing attack [3].

Previous research has shown that women are more likely to click on a malicious link in a phishing attack[15, 23]. A phishing email can use different techniques to influence people, these techniques could have a different effect on individuals. Younger people are more likely to get victimized when the phishing attack is a fake email from someone with authority[23]. For example, an email from the government, or a professor from the university. Another effective technique is creating scarcity, for example, an email that suggests

that a special offer will end soon. When compared to older people, the phishing email will be more effective when reciprocation is used. For example, clicking on a link for a coupon, while simultaneously malware will infect the computer of the victim [23].

2.3 Cyber security model

In this section, the CIA model will be explained in more detail.

2.3.1 CIA. This is a model to describe the information security attributes. Each of these properties should be as high as possible.

Confidentiality

Information should only be accessible to those people who are authorized to[27]. Especially in the case of the EHR, a lot of personal information is stored. This information is sensitive and should not be accessible by outsiders. A successful phishing attack has a direct impact on the confidentiality of the data because the attackers (who are unauthorized) have access into the system. To achieve high-level confidentiality, appropriate security measures should be implemented, for example, two-factor authentication or users could have different roles with different access to specific information.

Integrity

The second property of the CIA model is integrity, which means that the data should be accurate, complete,
and untampered[27]. Healthcare employees that work with the EHR, do need to rely on the data that is stored. The data should only be able to be modified by those people who have the right to do this. When comparing this to the confidentiality of an information system, not all people who have the right to access data should be able to modify it. A successful phishing attack has also a direct impact on the integrity of the information system, an attacker could potentially modify the data without permission.

Availability

Let's consider an information system with high levels of confidentiality and Integrity, designed to protect sensitive information from the EHR. However, the data is limited to access on a single computer per department in a hospital. The system is not useful at all in a hospital, because the data is not easily available for a healthcare employee. The third property of the CIA model is all about the availability of the data, the user should always have access to the information system when this is necessary. While a successful phishing attack may not directly impact the availability of the system, as the attacker only gains access into the system, it depends on the attacker's goal. When the attacker aims to decrypt all the data, ultimately the availability will be affected.

3 Method

3.1 Data collection

To answer the research questions, a systematic literature review and meta-analysis will be conducted. With the research questions, we would like to compare as many characteristics as possible, in the time frame of this thesis, this won't be possible with designing a new experiment. By using existing studies and performing a meta-analysis, this is possible, and more characteristics could be included.

Databases

To search for relevant articles and previous studies, electronic databases will be used. Those databases are Scopus, ACM Digital Library, and Sage journals.

Types of study

For this meta-analysis, studies that used experiments will be included. Experiments with phishing attacks are designed specifically for a study, which means that data is gathered with information about the characteristics of the email or the context that it is aligned with. This will give results that will show the effect of those characteristics.

First, as many studies as possible that are conducted within a healthcare context will be included, e.g. hospitals. When those studies won't be sufficient to perform an analysis, additional studies will be added to have a reliable dataset.

Search query

*((phishing OR "phishing attack" OR "phishing *mail") AND (effect OR influence OR characteristics) AND (persuasion OR influence OR technique) AND susceptibility AND (experiment* OR study OR studies OR "field trial")*

This search query is used to find studies that are in line with the types of studies that are described above; by including the search terms regarding an experiment, studies that are not designed like this, will most likely be excluded from the records.

Inclusion criteria:

- English language
- Dutch language
- Studies performed in a real-life situation, or simulated attacks in an organization
- Studies performed in a healthcare context

Exclusion criteria:

- Studies that are published before 2019
- Studies only focusing on theoretical frameworks, without empirical data
- Studies that are not accessible through the University of Twente
- Studies with a population that is too small, n <10

Data analysis

From the studies that are selected with the criteria mentioned above, one general dataset will be made. This data set will be analyzed with the statistical software R version 4.4.0 and R studio version 2024.04.1+748, with the "Metafor" package. The "ESC" package will be used when specific effect sizes should be converted to the odds ratio.

For each study, first will be tried to retrieve the raw data, which means that the sample size and the number of phished participants are used. For studies that do not provide these amounts, the effect size will be recalculated to an odds ratio with a 95% confidence interval. This can be calculated when a regression coefficient is provided.

$$OddsRatio = e^{\beta}$$

The expectation is that the studies will have a high level of heterogeneity, and because the populations are different, the random effect model will be used[1].

Overall phishing susceptibility

For all included studies that provide an overall sample size in combination with an overall click amount, susceptibility can be calculated in proportion. Finally, a meta-analysis of proportion will be performed on these studies. This will give us a good view in the overall susceptibility of a phishing attack.

Characteristics and effect on susceptibility

To include the right studies, two screening rounds will be conducted. The first round will aim to select studies that are performed as a real-life experiment, with empirical data. For the second screening round, the included articles will be screened again to determine if the data is valid and could be used in the general dataset. For the screening and general data set Microsoft Excel will be used. In this sheet, every data point will be added, related to one of the three research questions.

3.2 Screening

For the first screening one filter has been applied to all records, with this filter all the records that have been published before 2019 are excluded from the list. Cyber security is rapidly evolving from year to year, which means that older studies do not always represent the current state.

All the records that are published in or after 2019 are exported to an Excel document, where the title and DOI have been noted. Then, all the records are manually checked with the following criteria:

1. The record must be written in English or Dutch.
2. The study must be performed in a real-life situation, or simulated attacks in an organization with empirical data regarding cyber security susceptibility.
3. The study is an experiment that used a phishing email.
4. The study must measure the effect on susceptibility of human factors, elements from the phishing attack, or context-based factors.

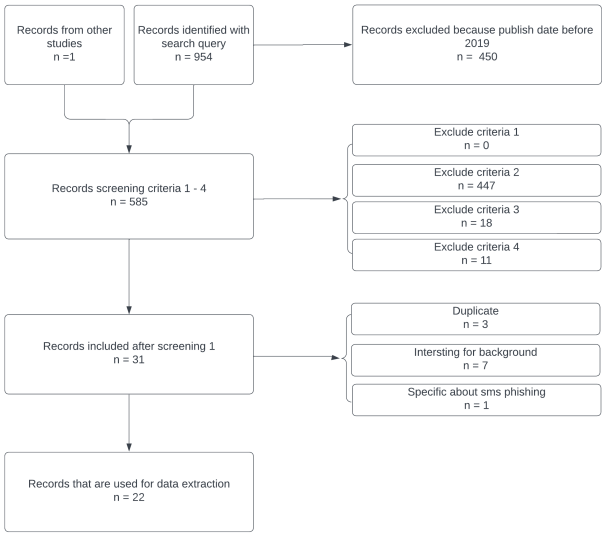


Figure 1. PRISMA flow

Each individual record has been screened on title and abstract by the researcher. In case of an unclear abstract where the goal of the study is vague, the full article has been read to determine if all the criteria have been met. In the Excel sheet, for every record, a YES in green has been noted if all criteria are met. For every record that does not meet a criterion, a NO has been noted with the corresponding criterion. In addition, a short explanation of why this record cannot be included is written for every article that is excluded.

This has been repeated with three different scientific databases: Scopus, ACM Digital Library, and Sage Journals. In Figure 1, a PRISMA flow diagram visualizes the screening process.

From all the records (n=955), 450 are excluded because of the publication date being before 2019. Of the 585 records that were published after or during 2019, 31 have been included. Finally, after a detailed look at each of the studies, 22 have been included for the data extraction. For each criterion, figure 1 shows how many records are excluded based on this criterion.

3.3 Extracting data

After the screening of all studies, the included articles are read in more detail. To get an overview of all the different studies, an Excel sheet has been made where all important information is noted. In this sheet, all factors that are included in the experiment have been categorized in one of the three research questions. This way, it is more convenient to see which characteristics are more commonly included in the

articles. All characteristics that have been included in the meta-analysis are explained in more detail:

Overall phishing susceptibility

First, a meta-analysis of proportion will be performed on the overall phishing rates of all studies that provide this data. Therefore, the total sample size of the study will be extracted, together with the total amount of clicks on the link in the phishing emails. This will give us the overall susceptibility to get phished with a phishing email.

Sex

RQ1 is about the characteristics of individuals, studies that provide data about the susceptibility between men and women are included in this meta-analysis. If raw data is provided (clicks and sample size), this will be used to calculate an odds ratio. In the case of another effect size, this will be converted to an odds ratio. Men are coded as 0 and women as 1.

Age

For the second characteristic of an individual, we will analyze if age influences the susceptibility of a phishing email. Therefore, we will categorize participants into 2 groups. The young population are participants younger than 40 years old. Where the participants in the old category are 40 years or older. This analysis will use the same method as mentioned for sex, either raw data will be calculated to the odds ratio, or other effect sizes will be converted to an odds ratio. Therefore, we coded the young group as 0 and the old group as 1.

Authority

Authority is a persuasion principle that is often used in a phishing email (RQ2). With this technique, an email will be designed with an authoritative tone, for example by imitating the government or using the name of someone's boss in the signature of the mail[8, 9, 18, 23]. Ultimately, the email will have the message that if the receiver won't comply, a sort of consequence will be taken, e.g. blocking an account or losing data. To determine if there is a significant effect of this persuasion technique, a sub-group analysis will be performed. For each study, the simulated phishing attack will be coded as using authority = 1, or not using authority = 0. Then, with the sample size and number of clicks, it is possible to calculate the proportion of the two groups.

Benefit

Another relevant persuasion technique is benefit, criminals can try to lure targets to click on a link to get something in return, for example, a free travel guide or a discount for an online shop[9, 18, 23, 35]. This technique could be used in many different forms, but the target will be tried to manipulate and click on the link in exchange for something. For benefit, there will also be performed a sub-group analysis.

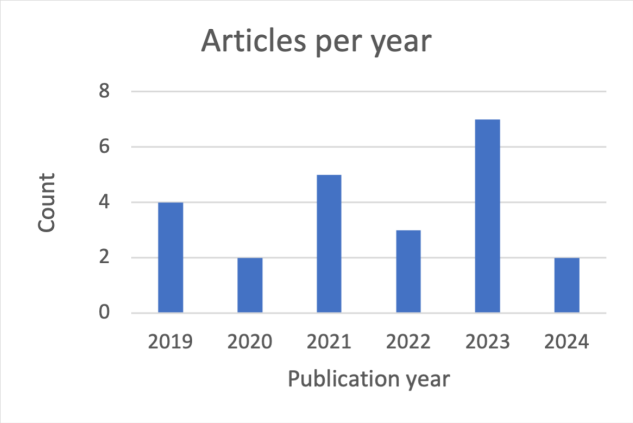


Figure 2. Articles per year

Liking

The last characteristic of RQ2 is the principle of using liking to persuade a target. This means that people are tending more willing to comply with people or things they like[9]. This persuasion principle can be used by cyber criminals, for example by sending fake Facebook messages, or using other elements that the target likes. To determine the effect of liking as a characteristic of the phishing email, a sub-group analysis will be performed.

Work-related

Healthcare employees could face work-related phishing emails. For example, an email about changing the password to maintain access into the systems. Previous research has shown that context alignment has a significant effect on the susceptibility of phishing attacks[5]. For the studies included, phishing emails that are aligned with the target's work context will be extracted and compared to those emails that are not. This will be done with an subgroup-analysis, the results will help us to develop an answer for RQ2.

Financial

Finally, phishing emails that are aligned with a financial context, for example, an email from a credit card company, will be included in the meta-analysis. This is a context that is often used in the included studies. Here, the same analysis will be performed as mentioned above for the work-related context.

4 Results

The 23 included studies are from 9 different countries, where most come from the United States, with n=10. From 1 study this is unknown. Table 1 provides an overview of the distribution of the articles and their year of publication. The most common publication year is 2023 (n=7), the distribution is shown in figure 2. Appendix A contains a table with specific information about every included study.

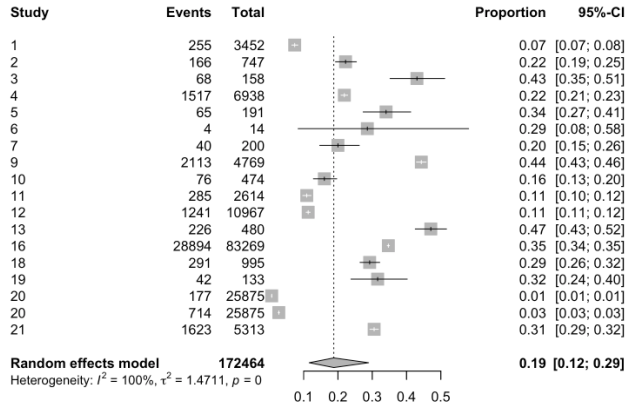


Figure 3. Overall phishing rate

4.1 Overall phishing susceptibility

For the overall phishing susceptibility, there is a total of n=172464 subjects, in the range of 14 to 83269 participants per study. For this analysis, there is a k=16 observations[6, 8, 10, 12, 14, 16, 23–26, 29, 30, 32–35] . When using the random effects model, there is a phishing susceptibility of 19% with a 95% confidence interval of [0.12;0.29]. Looking at the high level of heterogeneity with an I^2 being 100%, the random effect model is preferable for the studies that are used. This means that almost 1 out of 5 people will click on a link embedded in the phishing email. Figure 3 shows the forest plot for the overall phishing rate.

4.2 Characteristics of individuals

4.2.1 Sex. For sex, there is a total of n=11927 subjects, with 5234 men and 6068 women. In 2 studies, the descriptive research was not provided (n=607). There are k=8 observations in this analysis[2, 6, 7, 14, 16, 23, 24, 30]. For each study the odds ratio was already provided, or the effect size was converted to an odds ratio with an SE and confidence interval of 95%. For the calculations, the subjects were coded as 0=men and 1=women. The summary odds ratio is 0.93, which implies that women are less susceptible to a phishing email. However, with a p=0,127, this is not a significant effect as it exceeds the 0.05. The forest plot is shown in appendix B.

4.2.2 Age. For age there are k=3 observations. Of 2 observations the total participants were provided, which are 7069 together. With the third study not providing this data, the total subjects are unknown. Here, for the analysis young people were coded as 0 and old people as 1. The common odds ratio is 0.72, which indicates that people under 40 years old are more likely to fall for an phishing email. However, the p value is 0.4624, which means that this effect is not significant for the studies that are used in this analysis. The forest plot is visible in appendix B.

4.3 Characteristics of the email

4.3.1 Authority. For authority, there is a total of n=50950 subjects from k=12 observations[6–8, 10, 12, 14, 17, 26, 30, 33–35]. From those studies, there are 45 different groups where a phishing email is simulated, with a sample size between 14 and 6938 individuals. 14 of the simulated phishing attacks used the persuasion principle of authority, the other 31 did not use authority in the simulated phishing email. A subgroup analysis has been performed using the random effects model, because of the high level of heterogeneity with I^2 = 99.2%. The results are visible in Table 1. The forest plot in Appendix B gives an overview of all phishing attacks separated by authority.
However, there is no significant influence on the effect of the phishing email, because the difference between groups is Q=0.11 with p= 0.7457.

4.3.2 Benefit. For benefit, there is a total of n=50950 subjects from k=12 observations[6–8, 10, 12, 14, 17, 26, 30, 33–35]. Of the 45 different phishing attacks, 8 have used benefit as a persuasion principle. The other 37 did not use benefit in the email to manipulate the target. For benefit, the same subgroup analysis as before has been performed. The heterogeneity level is 99%, which indicates making use of the random effects model. The results are visible in Table 2. The forest plot in Appendix B gives an overview of all phishing attacks separated by benefit.
The overall p has a value of 0.7973, which means that there is no significant effect on the susceptibly of a phishing email that contains benefit as persuasion principle.

4.3.3 Liking. For Liking, there is a total of n=50950 subjects from k=12 observations[6–8, 10, 12, 14, 17, 26, 30, 33–35]. From all phishing attacks, 7 have used the persuasion principle of liking in the email. The other 38 did not use this technique. The heterogeneity is high with I^2 being 99%, so the random effects model has been used. The results are visible in Table 3. The forest plot in Appendix B gives an overview of all phishing attacks separated by Liking.
The overall p has a value of 0.0755, which indicates that there is no significant effect on the phishing susceptibility when liking is used as a persuasion principle in the email. However, the p-value is close to 0.05, which indicates that there could be an effect on the susceptibility but cannot be proven significant with this data set.

4.4 Context alignment of the email

4.4.1 Work context. For the subgroup analysis to determine the effect of receiving a phishing email aligned with a work context there is a total of n=58341 subjects from k=11 observations[6–8, 10, 12, 14, 17, 26, 30, 34, 35]. These 11 studies have performed a total of 35 simulated phishing attacks together. 18 attacks are aligned with the work context, the other 17 were not. The results of the subgroup analysis are

Table 1. Subgroup analysis authority

AUTHORITY USED	K	PROPORTION	95%-CI	TAU^2	TAU	Q	I^2 (%)
YES	14	0.1651	[0.1205; 0.2220]	0.3587	0.5989	107.3000	87.9
NO	31	0.1764	[0.1351; 0.2270]	0.7722	0.8787	5556.4200	99.5

Table 2. Subgroup analysis benefit

BENEFIT USED	K	PROPORTION	95%-CI	TAU^2	TAU	Q	I^2 (%)
YES	8	0.1615	[0.0867; 0.2810]	0.9782	0.9891	1614.0400	99.6
NO	37	0.1753	[0.1409; 0.2161]	0.5884	0.7671	3747.8000	99.0

Table 3. Subgroup analysis liking

LIKING USED	K	PROPORTION	95%-CI	TAU^2	TAU	Q	I^2 (%)
YES	7	0.2351	[0.1677; 0.3191]	0.2887	0.5373	471.9700	98.7
NO	38	0.1627	[0.1282; 0.2043]	0.7001	0.8367	4307.9200	99.1

visible in Table 4. Appendix B provides the forest plot.
The overall p has a value of 0.0816, which indicates that there is no significant effect of aligning the email with the work context. However, the p does not exceed 0.05 as much as the other factors. This could imply that there is an effect, but with this specific dataset not significant.

4.4.2 Financial context. Finally, for the subgroup analysis about the effect of using financial context alignment, there is a total of n=58341 subjects from k=11 observations[6–8, 10, 12, 14, 17, 26, 30, 34, 35]. From 35 simulated phishing attacks, 5 have been aligned with a financial context, the other 30 have not. The results are shown in Table 5, and the forest plot is shown in Appendix B.
We can see that the proportion of phishing susceptibility is lower among the group of participants that received a mail aligned with a work context. The p-value is 0,3483, which means that this effect on the susceptibility is not significant.

5 Conclusion

The outcomes of the meta-analysis provide valuable insights into the influence of various characteristics on the effect of phishing emails.

For the first research question, we analyzed individual characteristics to determine their impact on the effectiveness of a phishing email. The results indicate that there is no significant difference between men and women, to get victimized by a phishing email. The second individual character that was analyzed was age. The result suggests that there is no significant effect on the susceptibility between people that are younger than 40 years old, and people that are 40 years or older. Looking at the healthcare sector, more than 25% of healthcare workers are older than 55, which has only risen in recent years[17].

For the second research question, we analyzed what influence characteristics of phishing emails themselves have on the effect of the phishing email. Authority, benefit, and liking all do not have a significant influence on the effect of a phishing attack. Knowing that these principles do not influence the risk, does not mean that they cannot be used in phishing trainings. This study has shown that these principles are often used in phishing attacks, and thus should be considered to make people aware of this.

For the third research question, the effect of context alignment on phishing susceptibility was analyzed. We found that a phishing email that is aligned with the working context of the target, will increase the odds of getting victimized with 1.7, but this is not significant for the dataset that is used. In healthcare, email is often used for communication, which means that receiving a phishing attack aligned with this context could potentially lead to a data breach. However, due to the low number of studies that included context alignment in the experiment, this influence on the effect of phishing susceptibility is not very reliable. Also, the financial context does not have a significant effect on the phishing susceptibility.

Finally, the overall phishing rate stands at 19%. This implies that if an attacker sends a phishing email to 5 healthcare employees, approximately one of them will get victimized and potentially grant the attacker access to the Electronic Health Records.

Table 4. Subgroup analysis work context

WORK CONTEXT	K	PROPORTION	95%-CI	TAU^2	TAU	Q	I^2 (%)
YES	18	0.2198	[0.1620; 0.2912]	0.6084	0.7800	2031.9800	99.2
NO	17	0.1437	[0.0970; 0.2076]	0.8483	0.9210	3225.2300	99.5

Table 5. Subgroup analysis financial context

FINANCIAL CONTEXT	K	PROPORTION	95%-CI	TAU^2	TAU	Q	I^2 (%)
YES	5	0.1384	[0.0767; 0.2372]	0.5321	0.7295	453,1600	99.1
NO	30	0.1863	[0.1410; 0.2420]	0.8186	0.9048	5301.9300	99.5

6 Discussion

6.0.1 Limitations. The primary limitation of this meta-analysis is the low number of studies k on most characteristics. Papers usually do not provide the necessary data to perform a meta-analysis with binary data. This lack of detailed data is limiting the generalizability of the findings. Another reason why papers could not be included in specific meta-analyses on characteristics is the fact that there has not been a control group in the experiment.

A potential solution to this issue could have been to include studies from before 2019. However, due to Covid 19, the work situation has changed, and people more often work remoted from home. This shift in the work situation could influence phishing susceptibility. When looking at healthcare, this is not the case, as healthcare employees cannot work from home. This would mean that for healthcare, this shift is less relevant. For research question 3, it is noteworthy that fewer studies have been conducted on the effect of context on phishing susceptibility. For this study, the mail that is used is analyzed and determined if the mail uses a work-related context or a financial context. When neither of them is used, this is categorized as the control group. This way it was still possible to calculate odds ratios, but de validity is lower. Instead, experiments that clearly design an experiment with neutral emails and emails in a specific context, would have a higher validity. The fact that the studies have been coded by only one researcher is also a limitation of the validity. Finally, there is a lack of experiments conducted in a healthcare setting. This gap in the literature means that we still do not understand the phishing susceptibility in a healthcare context.

6.0.2 Future work. Future research should aim to conduct more experiments focusing on the specific characteristics identified in this study. It is essential that these studies include control groups that use neutral phishing emails, that do not contain persuasion principles or are designed for a specific context. This would improve the validity of results and provide a better dataset for future meta-analyses.

Furthermore, there is need to conduct more experiments in a healthcare setting, such as hospitals. Understanding phishing susceptibility in these environments is important because of the sensitive data that is stored in the EHR. Identifying and analyzing factors that could influence the phishing susceptibility, could lead to more specific, and thus more effective cyber security trainings.

By addressing these gaps and focusing on these areas, future research will provide a more comprehensive understanding of phishing susceptibility across different contexts, which can be used to design more effective cyber security trainings.

6.0.3 Practical implications. This paper can be used to design specific cyber security training. Unfortunately, there are no significant outcomes. However, the principles and context characteristics are still very relevant, as they are used in real phishing attacks from criminals. Effective training programs addressing phishing attacks will enhance the cybersecurity state of healthcare organizations, improving the Confidentiality, Integrity, and Availability (CIA) of their information systems.

References

[1] 2009. *Random-Effects Model.* 69–75. https://doi.org/10.1002/9780470743386.ch12

[2] Mohamad Alhaddad, Masnizah Mohd, Faizan Qamar, and Mohsin Imam. 2023. Study of Student Personality Trait on Spear-Phishing Susceptibility Behavior. *International Journal of Advanced Computer Science and Applications* 14, 5 (2023). https://doi.org/10.14569/ijacsa.2023.0140571

[3] Zainab Alkhalil, Chaminda Hewage, Liqaa Nawaf, and Imtiaz Khan. 2021. Phishing Attacks: A Recent Comprehensive Study and a New Anatomy. *Frontiers in Computer Science* 3 (2021). https://doi.org/10.3389/fcomp.2021.563060

[4] Shahryar Baki and Rakesh M. Verma. 2023. Sixteen Years of Phishing User Studies: What Have We Learned? *IEEE Transactions on Dependable and Secure Computing* 20, 2 (2023), 1200–1212. https://doi.org/10.1109/tdsc.2022.3151103

[5] Zinaida Benenson, Freya Gassmann, and Robert Landwirth. [n. d.]. Unpacking Spear Phishing Susceptibility. In *Financial Cryptography Workshops.*

[6] Marco De Bona and Federica Paci. 2020. A real world study on employees' susceptibility to phishing attacks. , Article 4 pages. https://doi.org/10.1145/3407023.3409179

[7] Roderic Broadhurst, Katie Skinner, Nick Sifniotis, and Bryan Matamoros-Macias. 2018. Cybercrime Risks in a University Student Community. *SSRN Electronic Journal* (2018). https://doi.org/10.2139/ssrn.3176319

[8] Pavlo Burda, Tzouliano Chotza, Luca Allodi, and Nicola Zannone. 2020. Testing the effectiveness of tailored phishing techniques in industry and academia: a field experiment. , Article 3 pages. https://doi.org/10.1145/3407023.3409178

[9] Robert B. Cialdini and Noah J. Goldstein. 2002. The Science and Practice of Persuasion. *Cornell Hotel and Restaurant Administration Quarterly* 43, 2 (2002), 40–50. https://doi.org/10.1177/001088040204300204 arXiv:https://doi.org/10.1177/001088040204300204

[10] Tom Cuchta, Brian Blackwood, Thomas R. Devine, Robert J. Niichel, Kristina M. Daniels, Caleb H. Lutjens, Sydney Maibach, and Ryan J. Stephenson. 2019. Human Risk Factors in Cybersecurity. , 87–92 pages. https://doi.org/10.1145/3349266.3351407

[11] T. Daengsi, P. Pornpongtechavanich, and P. Wuttidittachotti. 2022. Cybersecurity Awareness Enhancement: A Study of the Effects of Age and Gender of Thai Employees Associated with Phishing Attacks. *Educ Inf Technol (Dordr)* 27, 4 (2022), 4729–4752. https://doi.org/10.1007/s10639-021-10806-7

[12] Verena Distler. 2023. The Influence of Context on Response to Spear-Phishing Attacks: an In-Situ Deception Study. , Article 619 pages. https://doi.org/10.1145/3544548.3581170

[13] GOV.UK. 2020. Cyber Security Breaches Survey 2020. https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2020/cyber-security-breaches-survey-2020

[14] Frank L. Greitzer, Wanru Li, Kathryn B. Laskey, James Lee, and Justin Purl. 2021. Experimental Investigation of Technical and Human Factors Related to Phishing Susceptibility. *Trans. Soc. Comput.* 4, 2 (2021), Article 8. https://doi.org/10.1145/3461672

[15] Tzipora Halevi, Nasir Memon, and Oded Nov. 2015. Spear-Phishing in the Wild: A Real-World Study of Personality, Phishing Self-Efficacy and Vulnerability to Spear-Phishing Attacks. *SSRN Electronic Journal* (2015). https://doi.org/10.2139/ssrn.2544742

[16] Rand Abu Hammour, Yousef Al Gharaibeh, Malik Qasaimeh, and Raad S. Al-Qassas. 2019. The status of information security systems in banking sector from social engineering perspective. , Article 14 pages. https://doi.org/10.1145/3368691.3368705

[17] Marjolein J. Peters Hans Langenberg, Chantal Melser. 2023. Arbeidsmarktprofiel van zorg en welzijn in 2022. https://www.cbs.nl/nl-nl/longread/statistische-trends/2023/arbeidsmarktprofiel-van-zorg-en-welzijn-in-2022/3-zorgmedewerkers

[18] Michael Koddebusch. 2022. Exposing the Phish: The Effect of Persuasion Techniques in Phishing E-Mails. , 78–87 pages. https://doi.org/10.1145/3543434.3543476

[19] Katharina Krombholz, Heidelinde Hobel, Markus Huber, and Edgar Weippl. 2015. Advanced social engineering attacks. *Journal of Information Security and Applications* 22 (2015), 113–122. https://doi.org/10.1016/j.jisa.2014.09.005

[20] C. S. Kruse, B. Frederick, T. Jacobson, and D. K. Monticone. 2017. Cybersecurity in healthcare: A systematic review of modern threats and trends. *Technol Health Care* 25, 1 (2017), 1–10. https://doi.org/10.3233/THC-161263

[21] Elmer E. H. Lastdrager. 2014. Achieving a consensual definition of phishing based on a systematic review of the literature. *Crime Science* 3, 1 (2014), 9. https://doi.org/10.1186/s40163-014-0009-y

[22] Jane LeClair, Sherly Abraham, and Lifang Shih. 2013. An Interdisciplinary Approach to Educating an Effective Cyber Security Workforce. , 71–78 pages. https://doi.org/10.1145/2528908.2528923

[23] T. Lin, D. E. Capecci, D. M. Ellis, H. A. Rocha, S. Dommaraju, D. S. Oliveira, and N. C. Ebner. 2019. Susceptibility to Spear-Phishing Emails: Effects of Internet User Demographics and Email Content. *ACM Trans Comput Hum Interact* 26, 5 (2019). https://doi.org/10.1145/3336141

[24] David Maimon, C. Jordan Howell, Robert C. Perkins, Caitlyn N. Muniz, and Tamar Berenblum. 2021. A Routine Activities Approach to Evidence-Based Risk Assessment: Findings From Two Simulated Phishing Attacks. *Social Science Computer Review* 41, 1 (2021), 286–304. https://doi.org/10.1177/08944393211046339

[25] Gregor Petrič and Kai Roer. 2022. The impact of formal and informal organizational norms on susceptibility to phishing: Combining survey and field experiment data. *Telematics and Informatics* 67 (2022). https://doi.org/10.1016/j.tele.2021.101766

[26] F. Rizzoni, S. Magalini, A. Casaroli, P. Mari, M. Dixon, and L. Coventry. 2022. Phishing simulation exercise in a large hospital: A case study. *Digit Health* 8 (2022), 20552076221081716. https://doi.org/10.1177/20552076221081716

[27] C. Shah, D. Nachand, C. Wald, and P. H. Chen. 2023. Keeping Patient Data Secure in the Age of Radiology Artificial Intelligence: Cybersecurity Considerations and Future Directions. *J Am Coll Radiol* 20, 9 (2023), 828–835. https://doi.org/10.1016/j.jacr.2023.06.023

[28] Kevin F. Steinmetz and Thomas J. Holt. 2022. Falling for Social Engineering: A Qualitative Analysis of Social Engineering Policy Recommendations. *Social Science Computer Review* (2022). https://doi.org/10.1177/08944393221117501

[29] Chuan Tian, Matthew L. Jensen, Greg Bott, and Xin Luo. 2024. The influence of affective processing on phishing susceptibility. *European Journal of Information Systems* (2024), 1–15. https://doi.org/10.1080/0960085x.2024.2351442

[30] Chuan Tian, Matthew L. Jensen, and Alexandra Durcikova. 2023. Phishing susceptibility across industries: The differential impact of influence techniques. *Computers Security* 135 (2023). https://doi.org/10.1016/j.cose.2023.103487

[31] Zuoguang Wang, Limin Sun, and Hongsong Zhu. 2020. Defining Social Engineering in Cybersecurity. *IEEE Access* 8 (2020), 85094–85115. https://doi.org/10.1109/access.2020.2992807

[32] Ryan Wright, Steven Johnson, and Brent Kitchens. 2023. Phishing Susceptibility in Context: A Multilevel Information Processing Perspective on Deception Detection. *MIS Quarterly* 47, 2 (2023), 803–832. https://doi.org/10.25300/misq/2022/16625

[33] Rume Elizabeth Yoro, Fidelis Obukohwo Aghware, Bridget Ogheneovo Malasowe, Obinna Nwankwo, and Arnold Adimabua Ojugo. 2023. Assessing contributor features to phishing susceptibility amongst students of petroleum resources varsity in Nigeria. *International Journal of Electrical and Computer Engineering (IJECE)* 13, 2 (2023). https://doi.org/10.11591/ijece.v13i2.pp1922-1931

[34] Rume Elizabeth Yoro, Fidelis ObukohwoAghware, Maureen Ifeanyi Akazue, Ayei Egu Ibor, and Arnold Adimabua Ojugo. 2023. Evidence of personality traits on phishing attack menace among selected university undergraduates in Nigerian. *International Journal of Electrical and Computer Engineering (IJECE)* 13, 2 (2023). https://doi.org/10.11591/ijece.v13i2.pp1943-1953

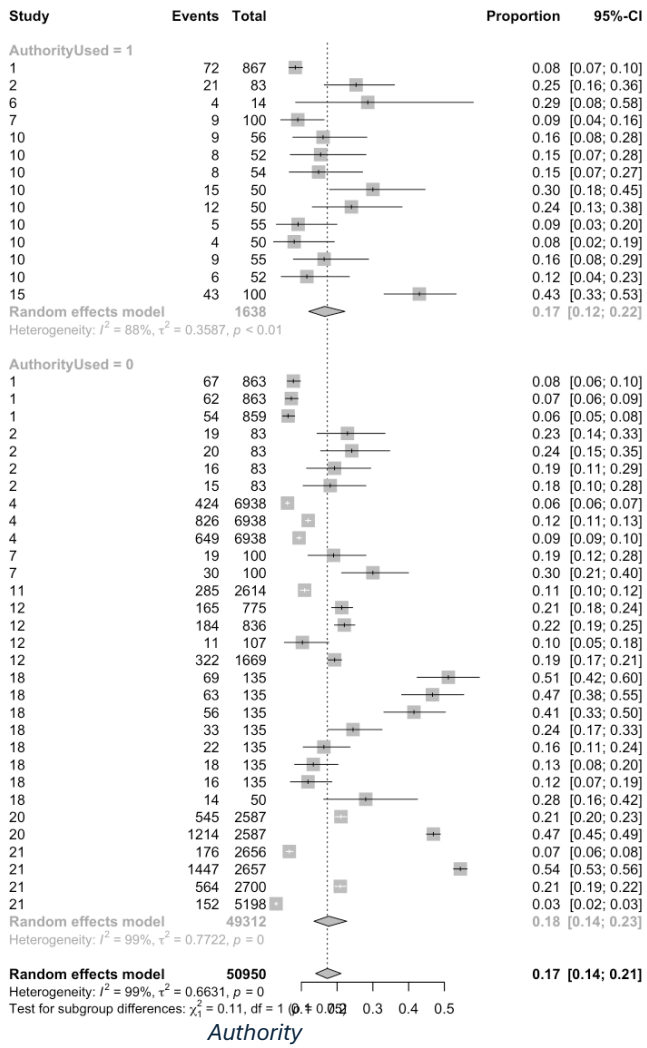
[35] Sijie Zhuo, Robert Biddle, Lucas Betts, Nalin Arachchilage, Yun Sing Koh, Danielle Lottridge, and Giovanni Russello. 2024. A Large-Scale Study of Device and Link Presentation in Email Phishing Susceptibility. , 78–85 pages. https://doi.org/10.1145/3638380.3638434

Appendix A

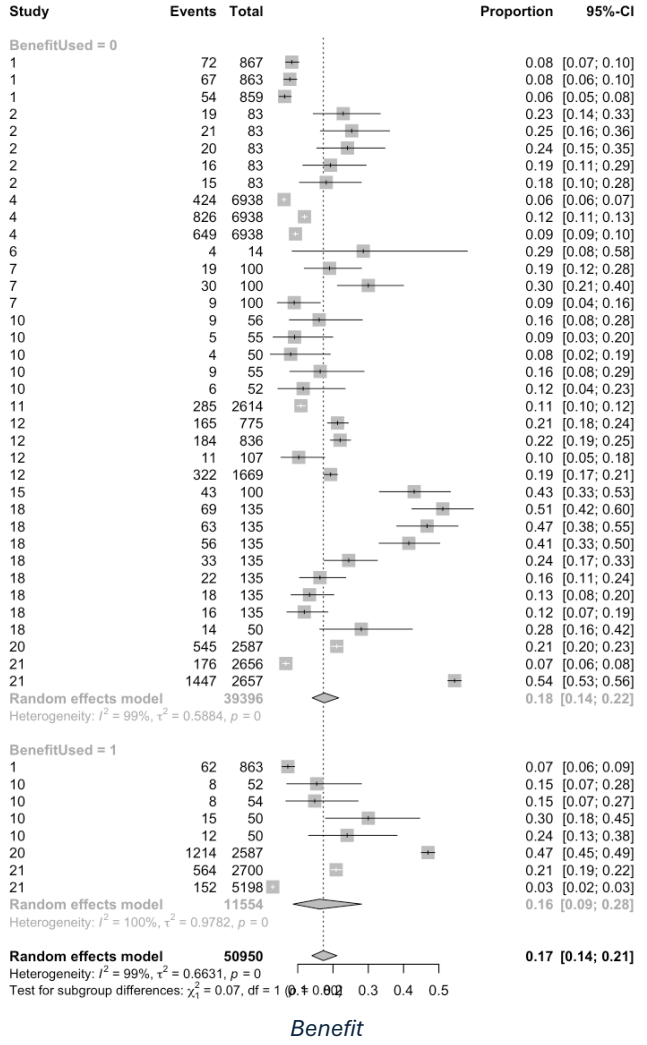
StudyID	Author	Title	DOI	Year	Country
1	M. Koddebusch	Exposing the Phish: The Effect of Persuasion Techniques in Phishing E-Mails	https://doi.org/10.1145/3543434.3543476	2022	Germany
2	P. Burda, T. Chotza, L. Allodi and N. Zannone	Testing the effectiveness of tailored phishing techniques in industry and academia: a field experiment	http://dx.doi.org/10.1145/3407023.3409178	2020	Netherlands
3	T. Lin, D. E. Capecci, D. M. Ellis, H. A. Rocha, S. Dommaraju, D. S. Oliveira, et al.	Susceptibility to Spear-Phishing Emails: Effects of Internet User Demographics and Email Content	https://doi.org/10.1145/3336141	2019	USA
4	F. L. Greitzer, W. Li, K. B. Laskey, J. Lee and J. Purl	Experimental Investigation of Demographic Factors Related to Phishing	https://doi.org/10.1145/3461672	2021	USA
5	M. D. Bona and F. Paci	A real world study on employees' susceptibility to phishing attacks	https://doi.org/10.1145/3407023.3409179	2020	Italy
6	V. Distler	The Influence of Context on Response to Spear-Phishing Attacks: an In-Situ Deception Study	https://doi.org/10.1145/3544548.3581170	2023	Germany
7	R. A. Hammour, Y. A. Gharaibeh, M. Qasaimeh and R. S. Al-Qassas	The status of information security systems in banking sector from social engineering perspective	https://doi.org/10.1145/3368691.3368705	2019	Jordan
8	M. L. Rahman, D. Timko, H. Wali and A. Neupane	Users Really Do Respond To Smishing	https://doi.org/10.1145/3577923.3583640	2023	USA
9	T. Cuchta, B. Blackwood, T. R. Devine, R. J. Niichel, K. M. Daniels, C. H. Lutjens, et al.	Human Risk Factors in Cybersecurity	https://doi.org/10.1145/3349266.3351407	2019	USA
10	C. Tian, M. L. Jensen, G. Bott and X. Luo	The influence of affective processing on phishing susceptibility	https://doi.org/10.1080/0960085X.2024.2351442	2024	USA
11	S. Zhuo, R. Biddle, L. Betts, N. Arachchilage, Y. S. Koh, D. Lottridge, et al.	A Large-Scale Study of Device and Link Presentation in Email Phishing Susceptibility	https://doi.org/10.1145/3638380.3638434	2024	Unknown
12	C. Tian, M. L. Jensen and A. Durcikova	Phishing susceptibility across industries: The differential impact of influence techniques	https://doi.org/10.1016/j.cose.2023.103487	2023	USA
13	R. E. Yoro, F. O. Aghware, B. O. Malasowe, O. Nwankwo and A. A. Ojugo	Assessing contributor features to phishing susceptibility amongst students of petroleum resources varsity in Nigeria	10.11591/ijece.v13i2.pp1922-1931	2023	Nigeria
14	R. E. Yoro, F. ObukohwoAghware, M. I. Akazue, A. E. Ibor and A. A. Ojugo	Evidence of personality traits on phishing attack menace among selected university undergraduates in Nigerian	10.11591/ijece.v13i2.pp1943-1953	2023	Australia
15	M. Alhaddad, M. Mohd, F. Qamar and M. Imam	Study of Student Personality Trait on Spear-Phishing Susceptibility Behavior	10.14569/IJACSA.2023.0140571	2023	Malaysia
16	G. Petrič and K. Roer	The impact of formal and informal organizational norms on susceptibility to phishing: Combining survey and field experiment data	10.1016/j.tele.2021.101766	2022	USA
17	P. Bayl-Smith, R. Taib, K. Yu and M. Wiggins	Response to a phishing attack: persuasion and protection motivation in an organizational context	10.1108/ICS-02-2021-0021	2021	Australia
18	R. Broadhurst, K. Skinner, N. Sifniotis	Phishing risks in a university student community	10.2139/ssrn.3176319	2019	Australia

	and B. Matamoros-Macias				
19	R. Wright, S. Johnson and B. Kitchens	Phishing Susceptibility in Context: A Multilevel Information Processing Perspective on Deception Detection	https://doi.org/10.25300/MISQ/2022/16625	2023	USA
20	D. Maimon, C. J. Howell, R. C. Perkins, C. N. Muniz and T. Berenblum	A Routine Activities Approach to Evidence-Based Risk Assessment: Findings From Two Simulated Phishing Attacks	https://doi.org/10.1177/08944393211046339	2021	Israel
21	F. Rizzoni, S. Magalini, A. Casaroli, P. Mari, M. Dixon and L. Coventry	Phishing simulation exercise in a large hospital: A case study	https://doi-org.ezproxy2.utwente.nl/10.1177/20552076221081716	2022	Italy
22	M. Canham, C. Posey, D. Strickland and M. Constantino	Phishing for Long Tails: Examining Organizational Repeat Clickers and Protective Stewards	https://doi-org.ezproxy2.utwente.nl/10.1177/2158244021990656	2021	USA
23	Z.Liu, L. Zhou, D. Zhang	Effects of Demographic Factors on Phishing Victimization in the Workplace	NA	2021	USA

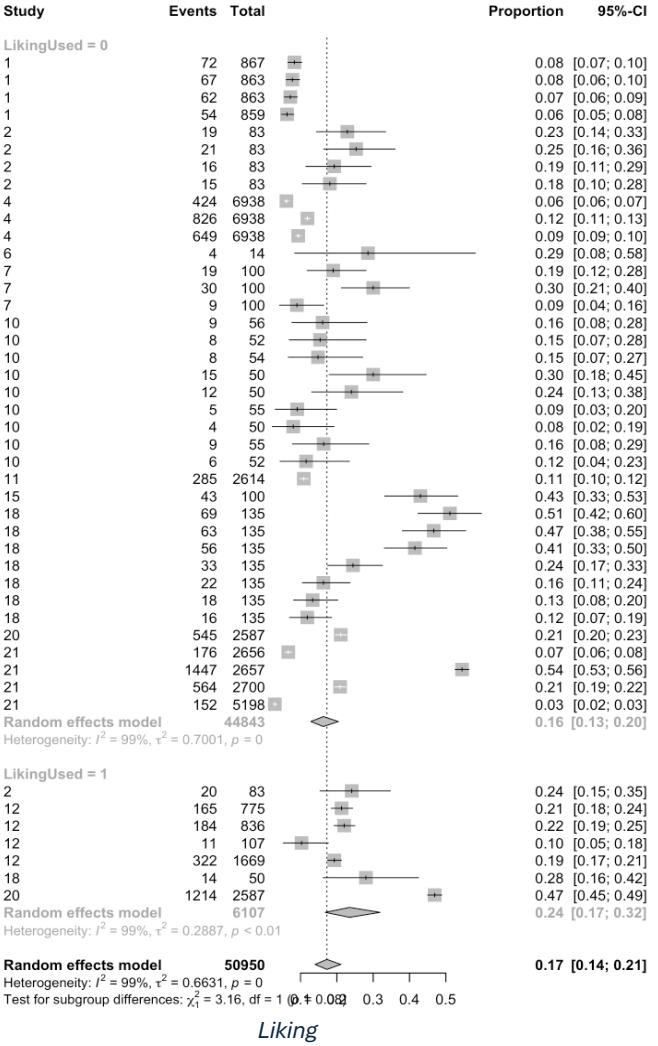
Appendix B Forest plots



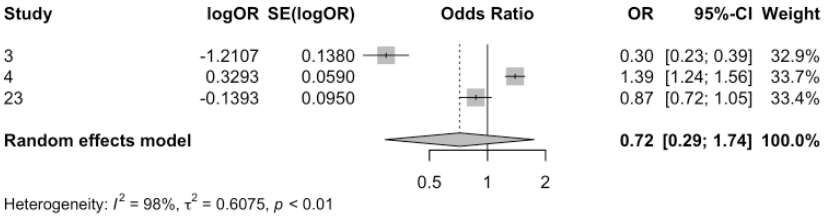
Authority



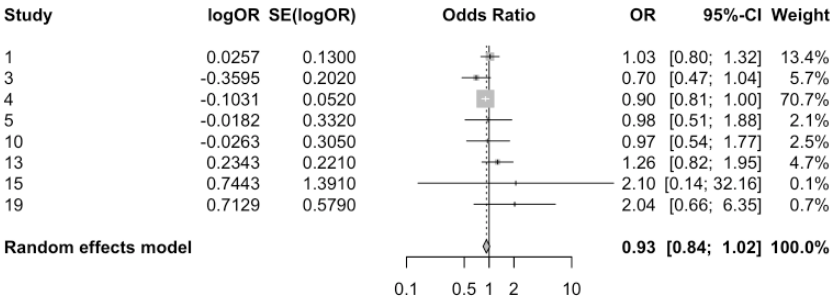
Benefit



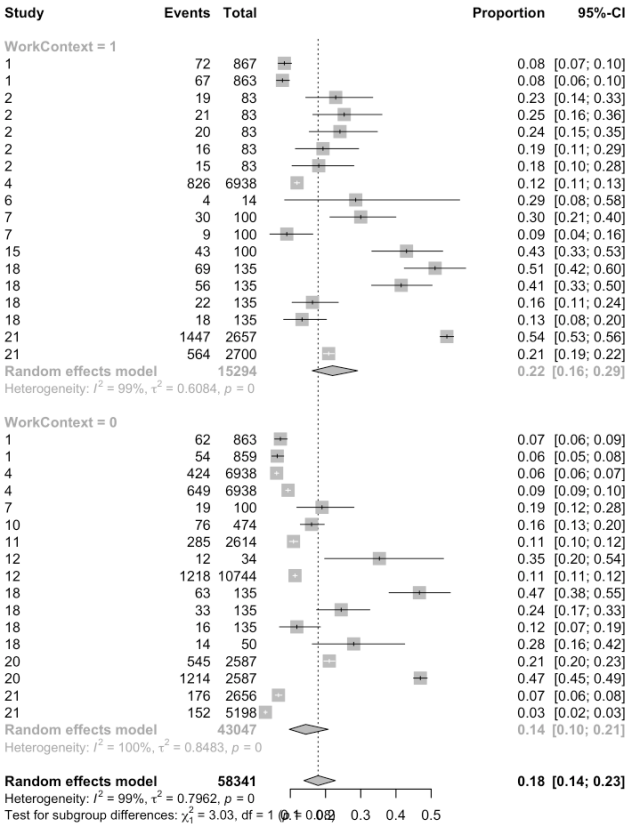
Liking



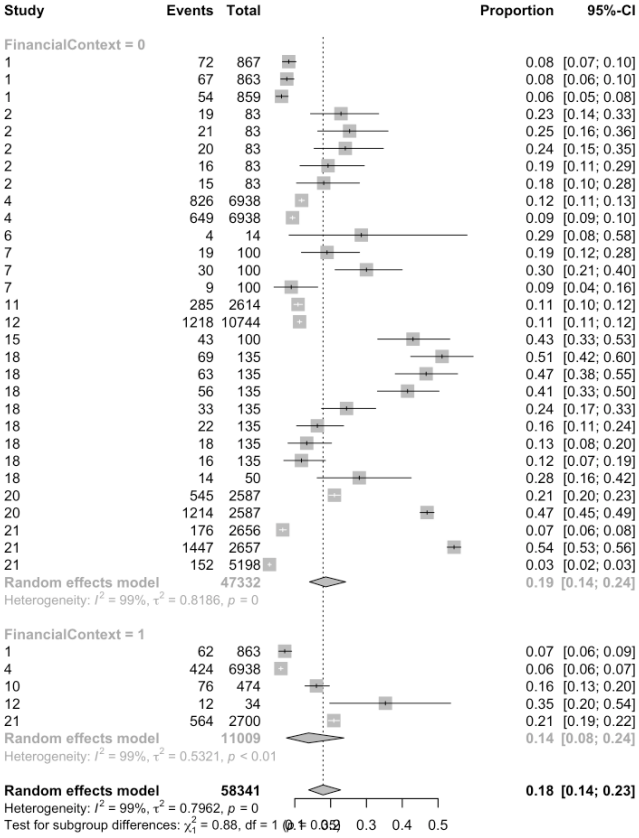
Age



Sex



Work context



Financial context