AI Integration in EHR-Based Pharmacovigilance: A Comparative Study of Germany and Egypt

David Naguib,[a] Georgi Chaltikyan,[b] Ahmed Adel[c]

[a]VigiPRO GmbH, Deggendorf Institute of Technology (DIT), Deggendorf, Germany
[b]Deggendorf Institute of Technology (DIT)/European Campus Rottal-Inn (ECRI), Pfarrkirchen, Germany
[c]Cur-is Life Sciences Solutions®, The University of Utah, Salt Lake City, Utah, USA

DOI: 10.25929/6q9v-cz19

ABSTRACT

Introduction: **Pharmacovigilance (PV) depends mainly on traditional reporting as a main source of data. This research will focus on another source, namely EHRs (electronic healthcare records). As we deal with big data from EHRs, AI (artificial intelligence) tools will be indispensable for the processing, and analysis of data and the early detection of ADRs (adverse drug reactions) from EHRs. In this research, we will explore the knowledge, attitudes and practices (KAPs) of the experts regarding the current application of AI in EHR-based PV, the potential benefits of implementing these technologies in PV, and the challenges toward their implementation in Germany and Egypt.**

Methodology: **A semi-structured survey of 30 questions that targeted the attitudes, knowledge and experience from PV experts (172 responses) was conducted.**

Results: **The results revealed that most PV companies in Egypt or Germany do not use EHRs as a main data source. This can be attributed to the lack of the application of EHRs in Germany and Egypt (e.g. EHRs in Germany is in a very early phase). Most of the PV organizations in both companies do not use AI as well in their PV activities. There is also a lack of proper adherence to data protection regulations in Egypt. However, the participants in both countries show a very positive attitude toward the adoption of AI and EHRs in the PV.**

Conclusion: **AI technologies and EHRS in the domain of PV are very rarely applied either separately or collectively in both countries, there is also a lack of knowledge among PV specialists about digital health but there are positive attitudes toward its adoption.**

KEYWORDS

Pharmacovigilance, electronic healthcare records, electronic healthcare records-based pharmacovigilance, adverse drug reactions, artificial intelligence

1. Introduction

One of the leading causes of death and a significant contributor to the expenses in public health (PH) are adverse drug reactions (ADRs) [1]. According to the Centers for Disease Control and Prevention (CDC), adverse medication events result every year in over 1.5 million visits to emergency rooms. About 500,000 individuals annually require hospitalization for further care following emergency room visits due to adverse medication reactions [2]. This accounts for the importance of pharmacovigilance (PV) which focuses on drug safety.
Pharmacovigilance can be defined as the science and process of monitoring the safety of medications and implementing actions to reduce the risks and increase the safety and benefits of medicines. This very task of the PV is one of the key roles of public health [3].

PV depends on a variety of data sources to monitor the safety of medications. Pharmaceutical companies, patients, and healthcare professionals voluntarily report ADRs through individual case safety reports (ICSRs) to spontaneous reporting systems (SRS) for compilation [4]. Furthermore, patient health data is recorded in EHRs during clinical practice and can be used to identify safety signals. Medical services are also tracked by healthcare claims databases, which helps with ADRs identification [5]. In addition, clinical trials, registries, and observational studies all shed light on drug use and safety in the real world and serve as potential and good supplemental sources of data for PV activities. Further viewpoints on adverse events and medication use can be found in patient surveys, pharmaceutical sales data, and social media conversations, in addition to the literature reviews (sometimes called medical literature) [6]. Collaborative networks (e. g. the European Network of Centers for Pharmacoepidemiology and Pharmacovigilance [ENCePP]) emphasize a complete strategy to guarantee patient well-being and drug efficacy, which adds to our understanding of medication safety [7]. So, there are a lot of sources of data that are utilized in the domain of PV.
In this research, the focus will be on EHRs (electronic healthcare records) as a promising source of data. The use of EHRs in pharmacovigilance is a good example of the secondary use of data [8, 9]. In the past ten years, there has been a rise in interest in pharmacovigilance using EHRs [10, 11]. And as we deal with the big data of the EHRs, we need advanced tools for data processing and analysis. So, the use of AI is crucial.

There is great interest in the topic of artificial intelligence (AI) and its application in the domain of PV. For instance, from 2010 to 2020 alone, the publication rate on the topic of machine learning (ML; a subcategory of AI) in PV has increased approximately sixfold [12].

EHRs are already utilized independently as a source of information for pharmacovigilance activities in some countries (e. g. South Korea). For example, the Korea Institute of Drug Safety and Risk Management (KIDS) transformed 9 million records into a certain common data model (CDM) to standardize the data in EHRs between different healthcare systems in order to facilitate the detection of drug safety signals [13]. Other European countries have already established EHR systems (Italy, Great Britain, Norway, Sweden, Finland, Denmark) [14]. Other countries like the U.S., Japan, Canada, and Australia have also already implemented EHRs and associated activities to the PV. The WHO also uses EHRs in the PV process [11, 15]. The use of EHRs in pharmacovigilance activities will be adopted as well by other international organizations (e. g. the Food and Drug Administration [FDA]) despite some challenges like the problem of heterogeneity and the integration between the drug safety data [16]. The FDA's Adverse Event Reporting System (FAERS) is working on getting Real-world Evidence (RWE) from real-world data (RWD) through a collaborative project called FDA's Sentinel Initiative. This project which is adopted by FAERS aims to use all the new technological advancements (big data, AI, EHRS, etc.) to enhance the monitoring of drug safety and this will be done through collaboration with different stakeholders (e. g institutes, hospitals, AWS, Epic developers, etc.) [16–18].

A comprehensive literature review of the existing literature on the use of AI in EHR-based pharmacovigilance in Germany and Egypt was implemented and there were no specific papers identified that discuss the use of AI or EHRs in Germany and Egypt separately or collectively. Almost all the

papers that address the pharmacovigilance situation in both countries are about common problems like the poor quality of the reports or the underreporting, [19] or German physician's lack of awareness about pharmacovigilance [20], or about the practices and attitudes of physicians and pharmacists in Egypt toward PV [21, 22]. Moreover, we identified case studies on some medications discussing the pharmacovigilance of homeopathic products or veterinary medicine or about the hierarchy and the organization of pharmacovigilance in Germany [23].

The research questions were aimed at detecting the attitudes, the knowledge and the behavior (KAPs) of the experts regarding the reality of AI application in EHR-based PV. Reality means the real applicability, namely: How is AI already being applied to EHR-based pharmacovigilance (EHRs as the mean source of data)? What are the advantages or benefits and opportunities of the future application of AI in EHR-based PV? And what are the challenges to such implementation of AI in EHR-based PV? For a better understanding of these dimensions, we need a real context, we therefore examined the reality, challenges, advantages, and opportunities of the applicability of EHR-based PV in both Germany and Egypt. Finally, we will pose some suggestions and recommendations that will help increase AI adoption in EHR-based PV in general, and in Egypt and Germany specifically.

To achieve these goals, an Internet-based survey was conducted by specialists in Germany and Egypt. Then all the results (either qualitative or quantitative data) were collected and analyzed to discover the common trends, opportunities, and challenges based on the expert opinions. Interviews with experts (8 experts) as a validation method for the semi-structured survey were also conducted.

The scope of this research is on the role of pharmacovigilance in post-market analysis. In this phase of post-market analysis, our focus will be on the detection of ADRs rather than the other functions of pharmacovigilance which are concerned with quality issues, adulteration, misuse…etc.
The survey followed this logical sequence, which is based on the research question. First, we examine the knowledge and practices regarding the reality and then the attitudes and the expectations of the experts regarding the potential opportunities, and the challenges toward the application, and finally the suggested solutions.

2. Methodology

2.1 Data collection

The **online-based survey** (quantitative/qualitative data collection techniques) is the research method which was used to collect the primary data, namely the attitudes, the knowledge and the opinions of the digital health specialists and healthcare professionals (HCP) who are involved in the process of pharmacovigilance in Germany and Egypt (target audience).
The Internet-based survey was conducted through a Microsoft Forms survey. The survey was started on 16th January 2024 and ended on the 22nd February 2024, and it was available in English and German.

2.2 Calculating the Sample of the Survey

To ensure the accuracy and representativeness of the data within the respective contexts of Germany and Egypt in this comparative study, distinct sample sizes were computed of each population separately. The study attempted to attain strong statistical power and precision within each group by calculating sample sizes appropriately, which would allow for relevant comparisons between Egypt and Germany. Sampling tactics ensured the inclusion of pharmacovigilance specialists as a crucial demographic to ensure proper representation, taking into consideration their smaller population when compared to other healthcare professionals.
A simple formula for determining sample size based on the population size and a desired confidence level was used.

The formula is:

$$\textbf{Sample size} = Z2\times p\times(1-p)/E2 \quad (1)$$

Where:
Z is the Z-score associated with the desired confidence level
p is the estimated proportion of the population that has the characteristic we were interested in (if unknown, we used 0.5 for maximum variability)
E is the margin of error

The desired sample sizes for Germany and Egypt were 60 and 90 participants, respectively. The difficulty in collecting bigger sample sizes was exacerbated by a scarcity of pharmacovigilance specialists in Egypt and Germany. The number of participants needed remained relatively constant despite the smaller population sizes (approximately 3,000) because the computation is mostly based on the acceptable margin of error and confidence level.

So, the confidence level was changed to roughly 90 % for **Germany**, where a smaller target sample size of 60 was established. Based on the assumption of 3,000 participants in pharmacovigilance, a proportion (p) of 0.5, and a margin of error (E) of 10.55 %, the revised sample size computation produced a result of around 60. Similarly, a comparable adjustment to almost 90 % for the confidence level was made for **Egypt**, with a target of 90 participants. Recalculating the sample size, we found that it was 90 with a population of 3,000 people participating in pharmacovigilance, a proportion (p) of 0.5, and a margin of error (E) of 8.57 %. These modifications made it possible to obtain more manageable sample sizes without sacrificing a respectable degree of accuracy and confidence in the survey results. These calculations were done with the help of online calculators [24] and ChatGPT.

In addition, the power of the experiment/statistical power was calculated for some of the data points (questions number 8, 10, 18). The one-sample z-test for proportions is the statistical test that is being applied in these situations for the validation or the confirmation of the hypothesis/claim for each of the questions of the survey. The mathematical formula for the z-statistic in a one-sample z-test for proportions is:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \quad (2)$$

One-sample Z test

$\hat{p}$ is the sample proportion.
p^0 is the hypothesized population proportion.
n is the sample size.

The above-mentioned equation was quoted from ChatGPT. The one-sample z-test for these questions was calculated for the general sample (total number of participants, Germany and Egypt) and each country separately. The Z for these experiments was conducted with the help of Python under the supervision and guidance of a data analyst.

The survey was designed to obtain the two main types of data, namely structured (labeled) and unstructured data (unlabeled), taking into consideration that our focus was on structured data [25].
To get this structured data, multiple-choice close-ended questions were used. Dichotomous or trichotomous (i. e., “yes” or “no”, “not sure”) questions were also employed in the survey to the same end. Rating or Likert scales were used as well [25].

2.3 Validation of the survey

Different ways were used to validate the survey [26].

Interviews: The opinions of the experts were collected through 8 interviews which ensured the validity of the content. The interviews were conducted as a validation method for the survey questions and to determine the best questions which serve the purpose of the survey. Interviews were conducted with different audiences working in different settings in PV fields (target audience), with different job descriptions. The interview was conducted within a time interval of 40–60 minutes. The interviews were conducted in English and German, recorded, and finally transcribed. Participants were asked to sign consent for the sake of data privacy and protection. The following table depicts the demographics of the 8 interviewees:

Table: Demographic information.

Age Categories
25–34: 2 interviewees
35–44: 5 interviewee
45–54: 1 interviewee
Over 65: no interviewees in this category
Gender
Male: 4 interviewees
Female: 4 interviewees
Workplace Categories
Pharmaceutical company: 5 interviewees
Contract research organization: 1 interviewee
Digital Pharmaceutical Company: 1 interviewee
Insurance company: 1 interviewee
Years of Experience
2–3 years: 3 interviewees
4-7 years: 3 interviewees
8+ years: 2 interviewees

PV role categories
Quality and PV manager: 2 interviewees
Senior pharmacovigilance specialist: 2 interviewees
Pharmacometrician: 1 interviewee
Pharmacist Product Owner and Manager: 1 interviewee
Pharmacovigilance physician | medical affairs: 1 interviewee
Patient tele-education: 1 interviewee
Country of Work
Egypt: 4 interviewees
Germany: 4 interviewees

Pilot test: a survey was done on a small scale (a sample of a similar target group of 10 people) before sending the survey to the whole target group to check the technical issues, the comprehensibility of the survey's questions, and the duration of the survey.

The questions of the surveys will be presented in the Results section together with the answers.
A list of 789 experts in Germany and Egypt that were to receive the survey was made (199 in Egypt, 590 in Germany).
The survey was sent through e-mails, social media platforms (e. g., specialized Facebook groups), and recruiters' platforms (LinkedIn), and it was further forwarded to personal contacts through colleagues.

3. Data Analysis

The data derived from the survey were divided into two categories: structured categorical, and unstructured data.

3.1 Analysis of the Structured Data

Regarding the structured categorical data, it was first saved in tabular format and rendered into quantitative form. The process of conversion of structured categorical data to structured quantitative data was done by using the measures of center (i. e., frequency or mood). These conversions were done by Microsoft Forms Analytics itself, and the data was collected in xlsx format and saved in One Drive. **Descriptive analytics** were conducted from Microsoft Forms data analytics and the data was also visualized using graphics (histogram, pie chart, dot plot).

The tabular data in the Microsoft Excel sheet format (xlsx) was imported to be analyzed by the Python programming language for further significant analytics to draw a comparison between Germany and Egypt. The comparison was only implemented for the questions which may show significant insights because of the comparison between Egypt and Germany. **These were the questions 9, 10, 13, 16, 18, 28, 29.**

The validation of this comparative analysis was conducted using **the two-sample z-test.** In our specific case, null hypothesis (H0): The percentages of affirmative answers from Egypt and Germany are the

same while the alternative hypothesis (H1) represents the percentages of favorable answers from Germany and Egypt differ.
The following mathematical formula of the two-sample z-test was employed:

$$z = \frac{(p_1 - p_2)}{\sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}} \quad (3)$$

Two-sample z-test

Where:
The sample proportions of the two groups under comparison are denoted by p1 and p2.
The sample sizes for the two groups are n1 and n2.
The total proportion combined from the two groups is denoted by p.
The two-sample z-test equation was quoted from ChatGPT.
The comparison and the validation were done using **Python** and the results were incorporated in the Results section.

All the graphs which were generated by Microsoft Forms or by Python. They were subsequently refined and modified (only for the better appearance) using ChatGPT.

3.2 Analysis of the Unstructured Data

The unstructured data from the survey was collected carefully to be analyzed by **thematic analysis** and incorporated into the Conclusion part.

4. Results

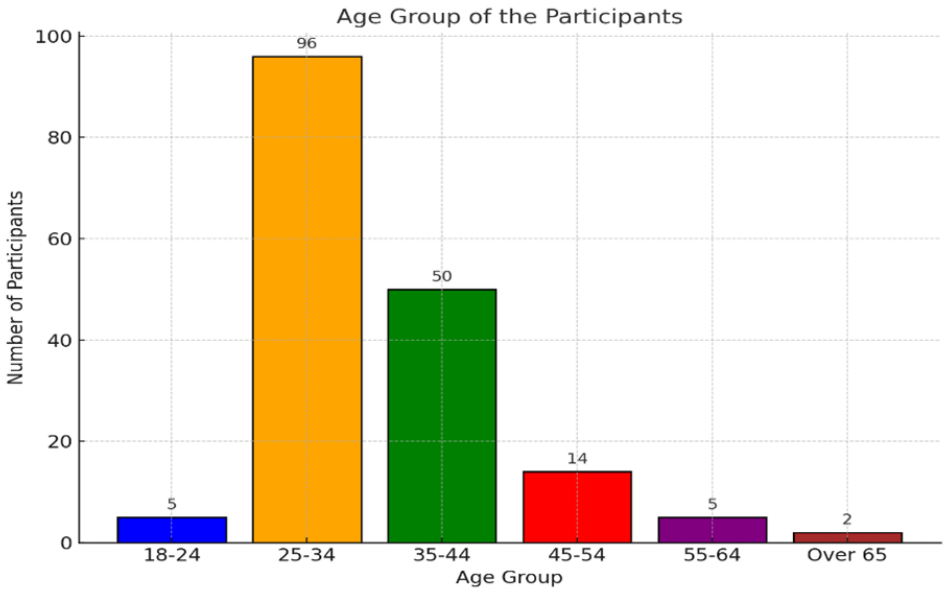
In total, **172 responses** to the survey were collected (63 in Germany, 93 in Egypt, and 16 other countries). The survey consisted of **30 questions which** were classified into 4 categories: q1–q7 consent of participation and demographics, q8–q19 Reality and application, q20–q23 (opportunities), and q24–q30 (challenges and suggestions).
The graphs represent the total number of participants including those who were not from Germany or Egypt to get a holistic overview of the topic of interest and because some of those who answered with “other” had already worked in Egypt or Germany
In addition, in certain questions where the statistical comparisons were deemed to be proper between Germany and Egypt, only data from both countries (excluding other participants) were utilized to conduct statistical analysis.
The results of one-sample z-test and the two-sample z-test were represented in association with corresponding questions.

4.1 Consent and Basic Demographics.

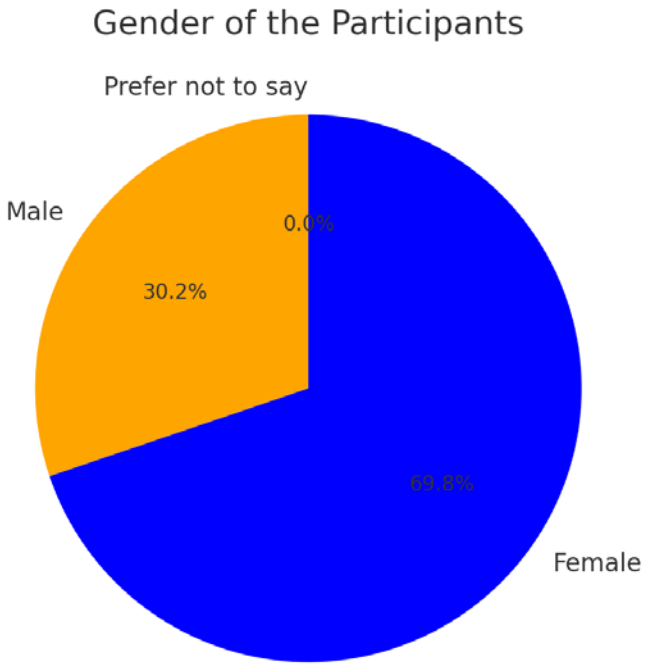
1. I have read and understood the information provided and have had the opportunity to ask questions. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving a reason and without cost.

I agree to participate: 172 I don't agree: 5

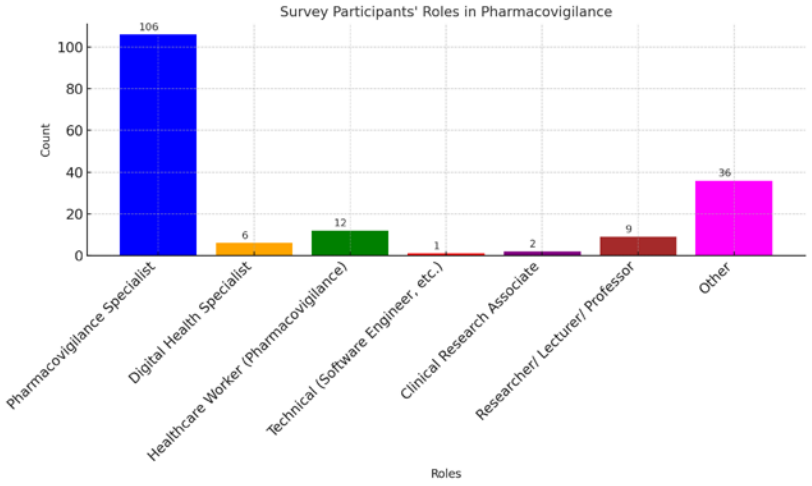
2. What is your age?



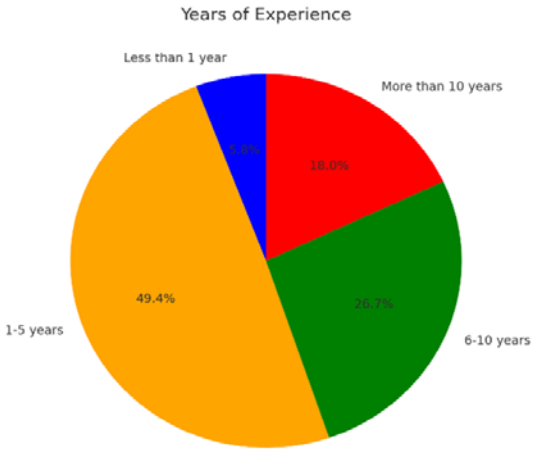
3. What is your gender?



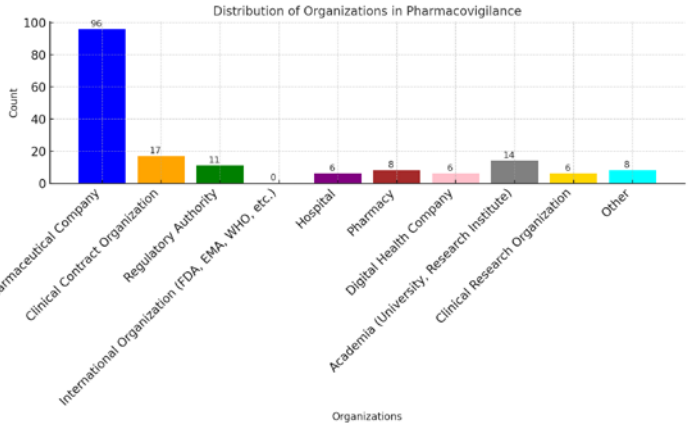
4. What is your current role and area of expertise in healthcare?



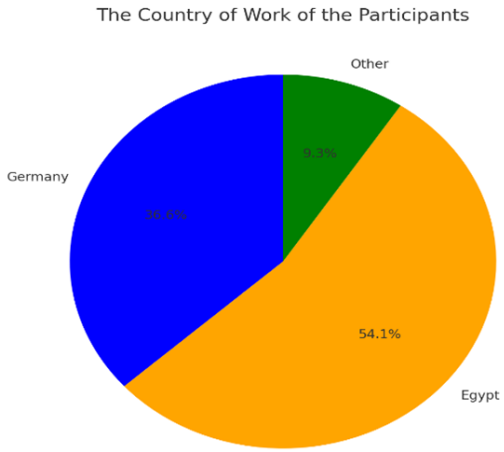
5. How many years of experience do you have in your current role?



6. Where is your workplace?

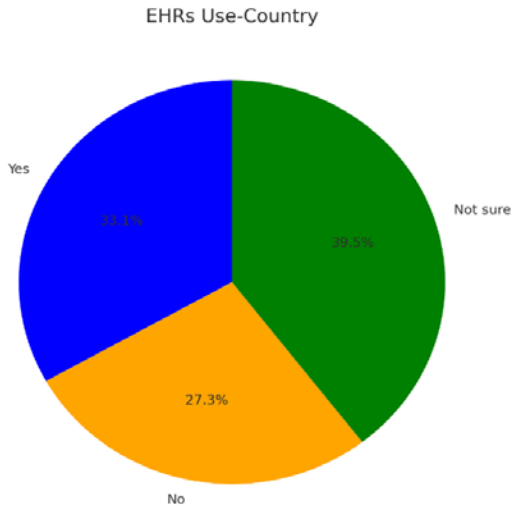


7. In which country are you currently working?



4.2 Reality and Application

8. Does your country of residence use electronic healthcare records (EHRs) in its healthcare system?



The one-sample z-test was conducted on this question and the results were the following:

For the total sample (Germany and Egypt)

The power of the experiment was 1.0; the z-statistic was 6.44704865098455, while the p-value was 1.140491864235056e–10. The p-value was 1.140491864235056e-10 less than the chosen significance level of 0.05. Therefore, we rejected the null hypothesis. There was sufficient evidence to suggest that the observed proportion of 0.3314 is significantly different from the expected proportion of 0.1.

For Germany only

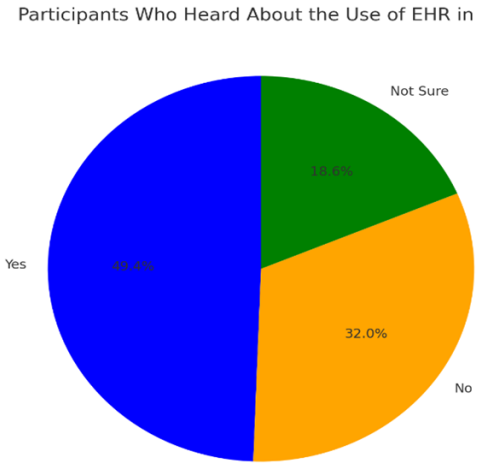
The power of the experiment was 1.0, the z-statistic 7.561249896677136, while the p-value was 3.992144237751295e–14. The p-value was 3.992144237751295e–14 less than the chosen significance

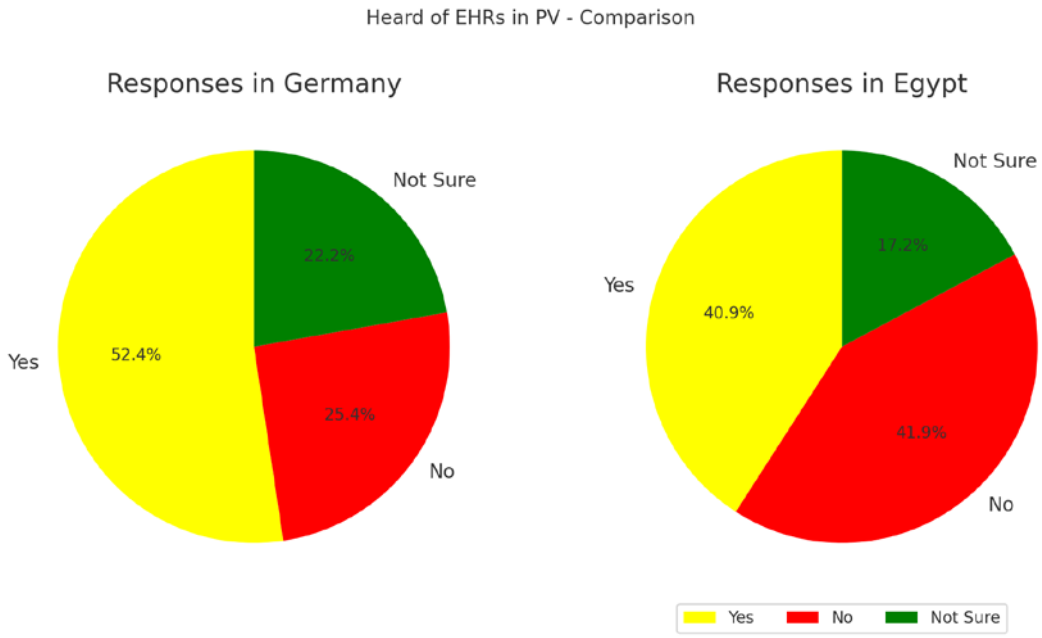
level of 0.05. Therefore, we rejected the null hypothesis. There was sufficient evidence to suggest that the observed proportion of 0.5714 is significantly different from the expected proportion of 0.1.

For Egypt only

The power of the experiment was 0.21639975207702283, while the p-value was approximately 0.269; it exceeds our chosen significance level of 0.05. Hence, we had no grounds to reject the null hypothesis. In other words, there was not enough statistical evidence to support the claim that the observed proportion of 0.1398 deviated significantly from its hypothesized value of 0.1.

9. Have you ever heard about the use of electronic healthcare records (EHRs) of patients in the pharmacovigilance process?

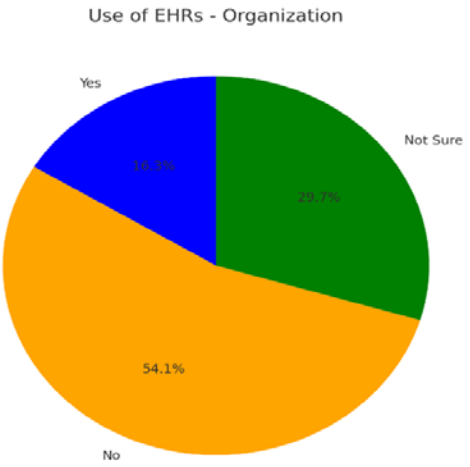


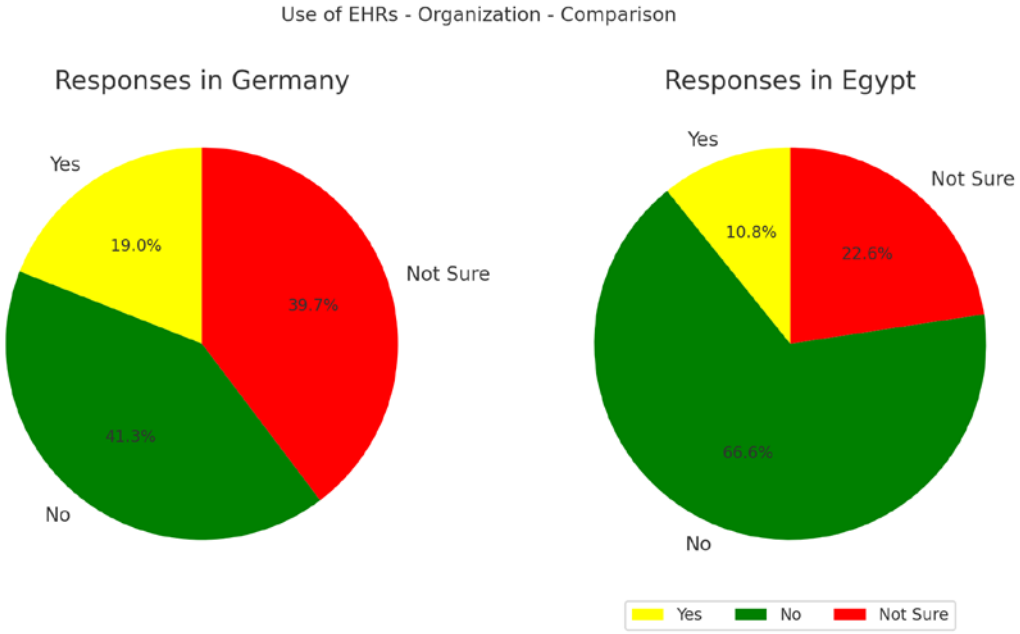


The **two-sample z-test** was conducted for this question.

The experiment had a power of 0.294467363572484. There were 93 samples in total from Egypt, which had a "yes" response in 38 cases and a "no or not sure" response in 55. There were 63 samples from Germany. Of them, only 33 provided "yes" responses, while the rest registered negative responses. The p-value for this research was calculated as approximately 0.156248483680624, with the z-statistic being approximately equal to that value at around 1.4178023301850076. As such, the null hypothesis remained unrefuted by the data used. There is no significant difference between the two countries.

10. Does your organization use Electronic Health Records (EHRs) in the pharmacovigilance process?





The **one-sample z-test** was conducted on this question and the results were the following:

Total sample (Germany and Egypt)

The experiment's power was 0.8063293867124136, the z-statistic was -3.0980984951602624, while the p-value was 0.0019476668523687802. Therefore, the null hypothesis was rejected since there was enough evidence to demonstrate that the observed proportion of 0.1628 differed significantly from the expected one of 0.25.

Germany Only

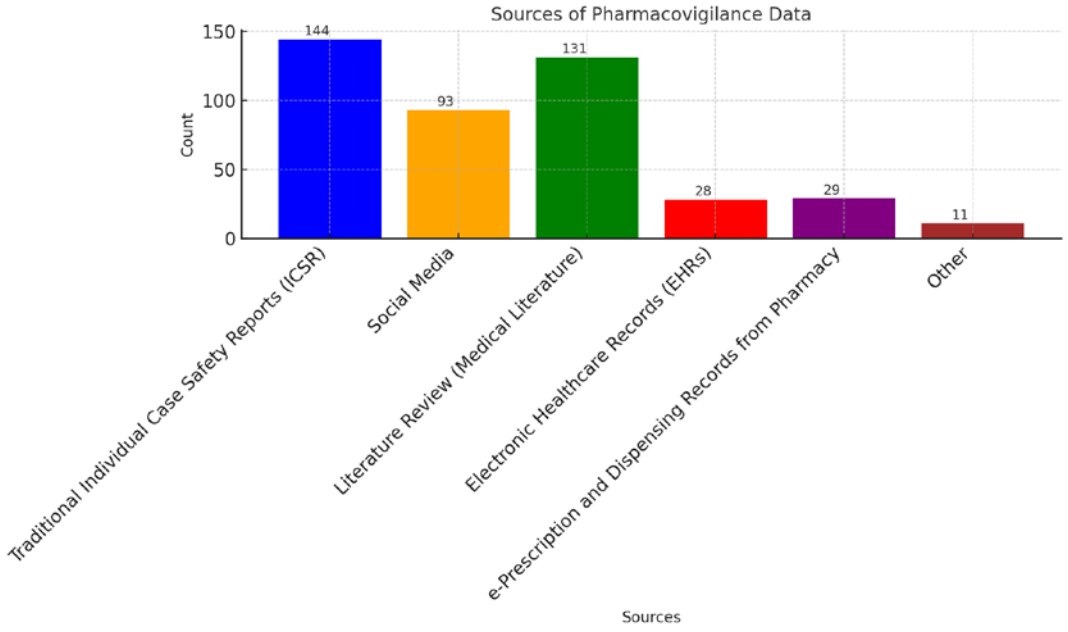
The experiment's power was 0.2028228819391893, the z-statistic was -1.2031667773495276, while the p-value was 3.992144237751295e-14. Therefore, we did not reject the null hypothesis since there was not enough evidence to show that the observed rate of (19 %) was different from the expected (25 %).

Egypt Only

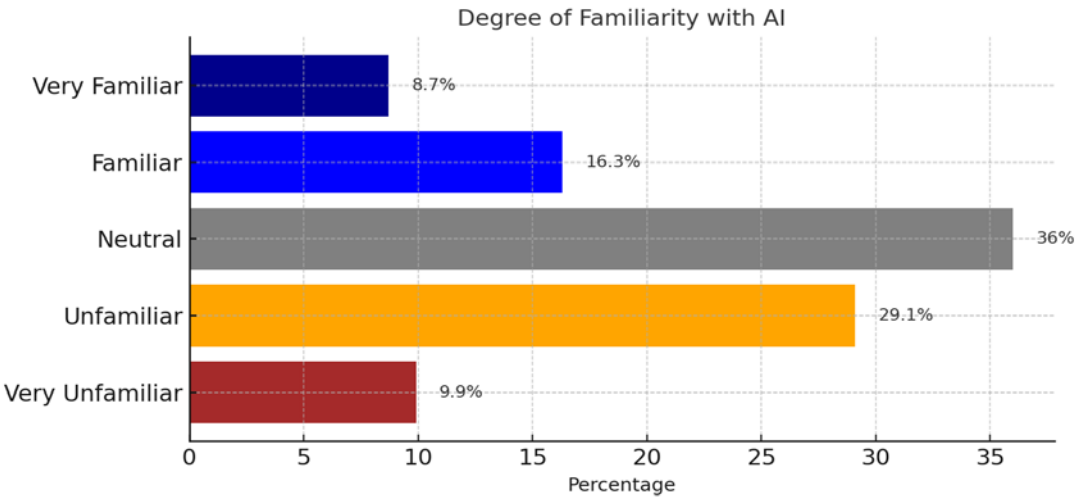
The power of the experiment was 0.9512066280912967 whereas the p-value was approximated as equal or greater than or equal to 0.05. Therefore, the null hypothesis should be rejected. There was enough evidence that demonstrated a significant difference between an observed proportion of 0.1075 and an expected one equaling 0.25

This question was subjected to a **two-sample z-test**. The power of the experiment was 0.3021145262608525. The z-statistic value was equal to 1.4605723670157307; while p = 0.14413283425964837.
The p-value equals approximately 0.144. This value was greater than or highly close enough to the pre-chosen significance level of $\alpha = 0.05$; therefore, we could not reject the null hypothesis here. There seemed to be no compelling evidence that the proportion of positive responses differed significantly between Germany (0.19) and Egypt (0.108).

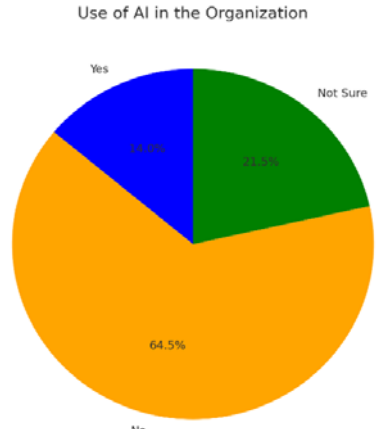
11. Which type of/source of information is used in the pharmacovigilance process in your workplace (select all that apply)?

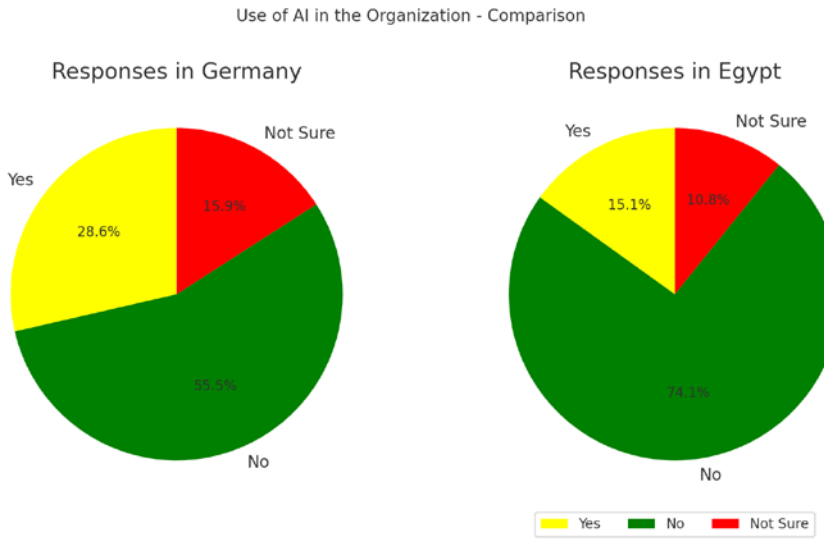


12. How familiar are you with the concept of AI (artificial intelligence) in pharmacovigilance?



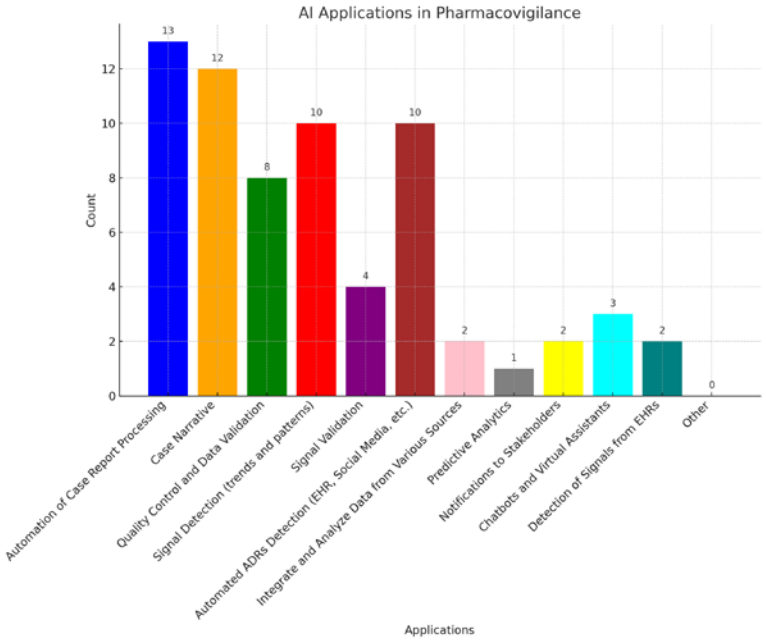
13. Does your organization currently use AI (artificial intelligence) in the process of pharmacovigilance?



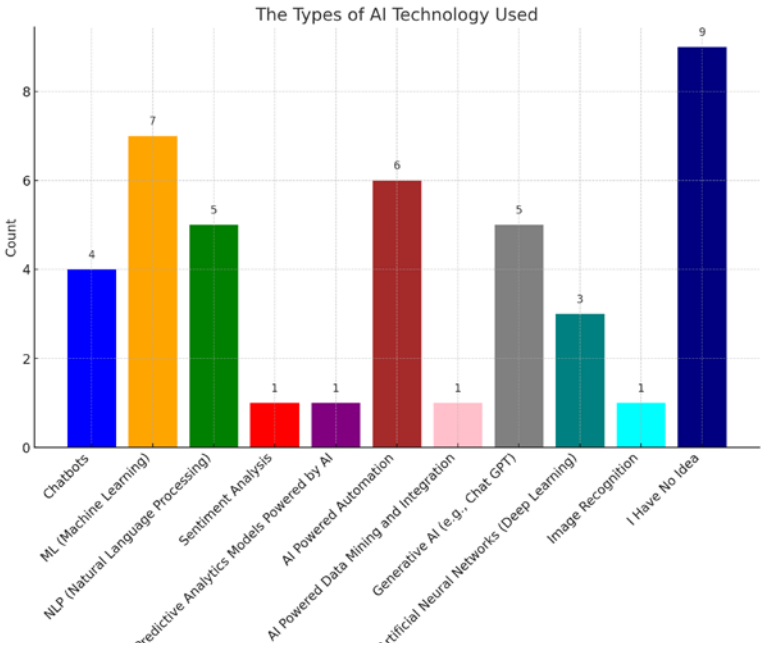


For this question, the two-sample z-test was conducted. The power of the experiment was 0.1529. The z-statistics were 0.9386 while the p-value was 0.348. The p-value was approximately 0.348, therefore, we failed to reject the null hypothesis. There was not sufficient evidence to suggest that there was a significant difference between the proportion of positive responses from Germany (0.159) and Egypt (0.108).

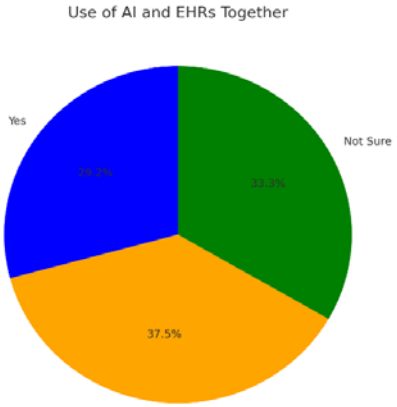
14. In which field does your organization use AI (artificial intelligence) in the pharmacovigilance process (select all that apply)?

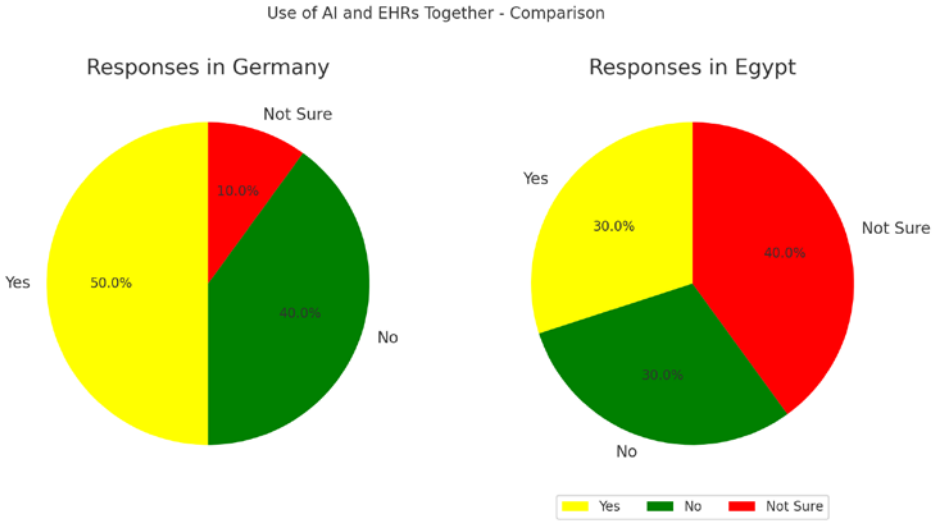


15. What specific AI (artificial intelligence) technologies or tools does your organization use for pharmacovigilance (select all that apply)?



16. Does your organization have any previous experience with the use of AI (artificial intelligence) and EHRs (electronic healthcare records) together, in other words, the use of AI technology in the detection of adverse drug reactions from the electronic healthcare records of the patients?



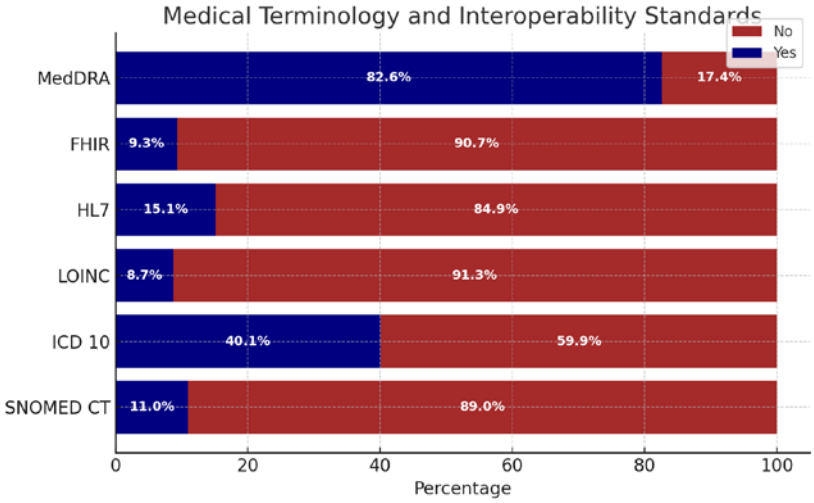


For this question, **the two-sample z-test** was conducted.

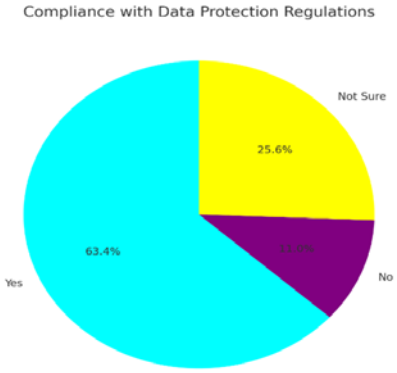
The power of the experiment was 0.21085961971387923. The z-statistic was -1.1180339887498947, while the p-value was 0.2635524772829728.

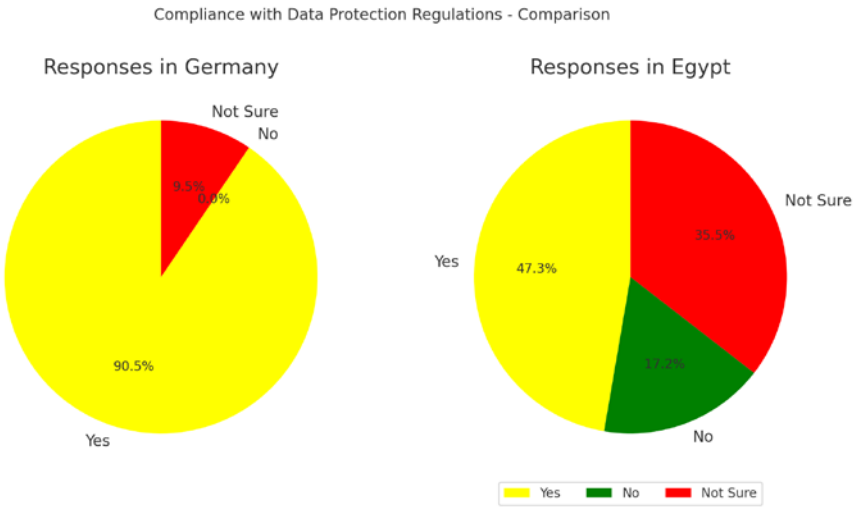
The p-value was approximately 0.264, therefore, we could not reject the null hypothesis.
There was not sufficient evidence to suggest that there was a significant difference between the positive responses from Germany (0.1) and Egypt (0.3).

17. Do you have any knowledge about the following concepts of medical terminology, ontology, data exchange, and interoperability?



18. Does your organization ensure compliance with the German DSGVO/GDPR (General Data Protection Regulations) or Egypt's data protection law?





The **one-sample z-test** was conducted on this question and the results were the following:

For the Total Sample (Germany and Egypt)

The power of the experiment was 0.9094870575812052. The z-statistic was -3.1652695177265 and the p-value was 0.0015493935939829518. The p-value is 0.0015493935939829518, therefore, we rejected the null hypothesis. There was sufficient evidence to suggest that the observed proportion of 0.6337 is significantly different from the expected proportion of 0.75.

For Germany Only

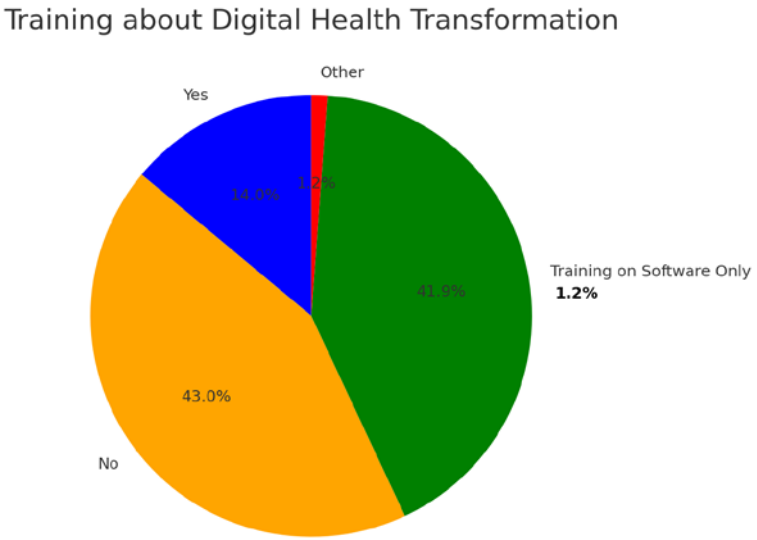
The power of the experiment was 0.9065155381549374. The z-statistic was 4.184675991984893, while the p-value was 2.8557312958510958e-05. Alternative: As the p-value was above the threshold of 0.05, we rejected the null hypothesis. There was sufficient evidence to suggest that the observed proportion of 0.9048 was significantly different from the expected proportion of 0.75.

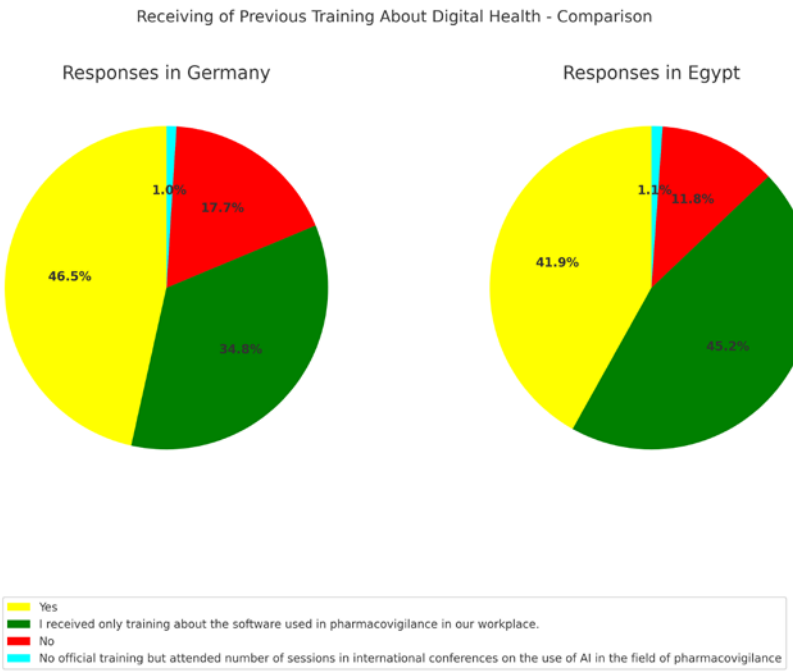
For Egypt Only

The power of the experiment was 0.9998068220766971, the p-value was approximately 0.0. Therefore, we rejected the null hypothesis. There was sufficient evidence to suggest that the observed proportion of 0.4731 was significantly different from the expected proportion of 0.75.

For this question, **the two-sample z-test** was conducted:
The power of the experiment was 0.9999834400364302.
The z-statistic was 5.536781025086471, while the p-value was 3.080813360109637e-08.
The p-value was approximately 0.0, which is below the chosen significance level of 0.05.
Therefore, we rejected the null hypothesis. There was sufficient evidence to suggest that there is a significant difference between the proportion of positive responses from Germany (0.905) and Egypt (0.473). The proportion of "yes" responses from Germany was higher.

19. Did you receive any formal training in digital health transformation?
Digital Health Transformation means briefly: the application of information and communication technologies in the healthcare and medical fields. It is also about the implementation of AI, big data, mobile apps, data-sharing principles, virtual and augmented reality, etc. in the healthcare sector.

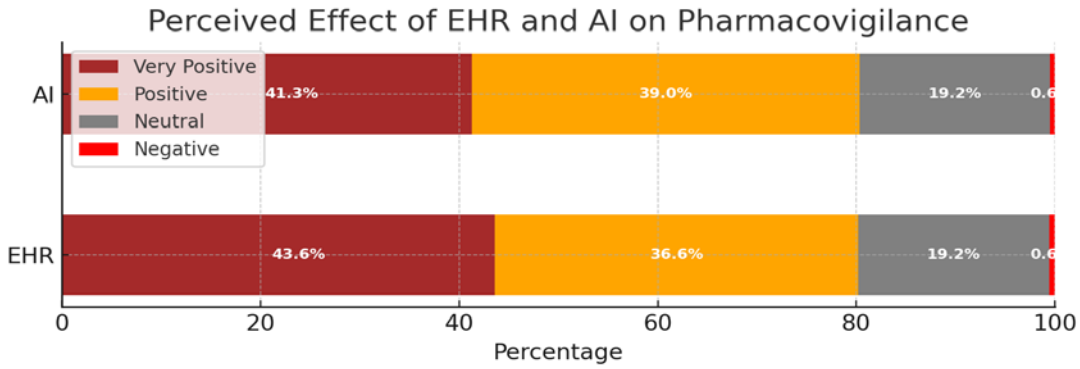




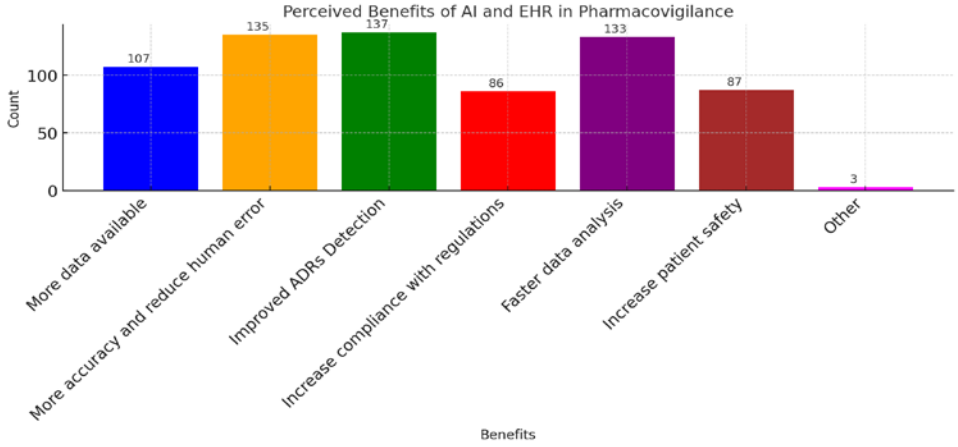
The two-sample z-test was conducted for this question.
The power of the experiment was 0.1653179334753394. The z-statistics were 0.9917466689612985, while the p-value was 0.3213211190920956.
The p-value was approximately 0.321. Therefore, we failed to reject the null hypothesis. There is not sufficient evidence to suggest that there is a significant difference between the proportion of positive responses from Germany (0.175) and Egypt (0.118).

4.3 Section: Opportunities

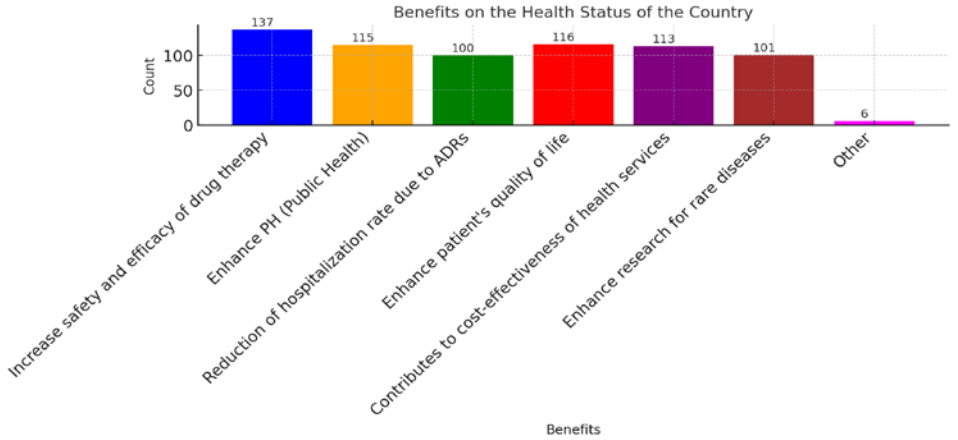
20. How do you see the effect of the following concepts, if they were applied, on pharmacovigilance:



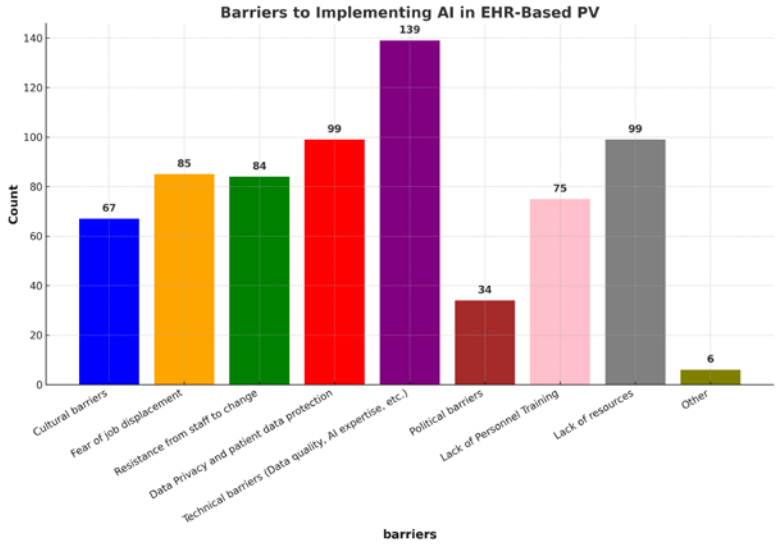
21. What are the expected benefits of using AI and EHR together in PV (select the most significant answers)?



22. What are the expected benefits of using AI in EHR-based pharmacovigilance on the health status of the country (select the most significant answers)?

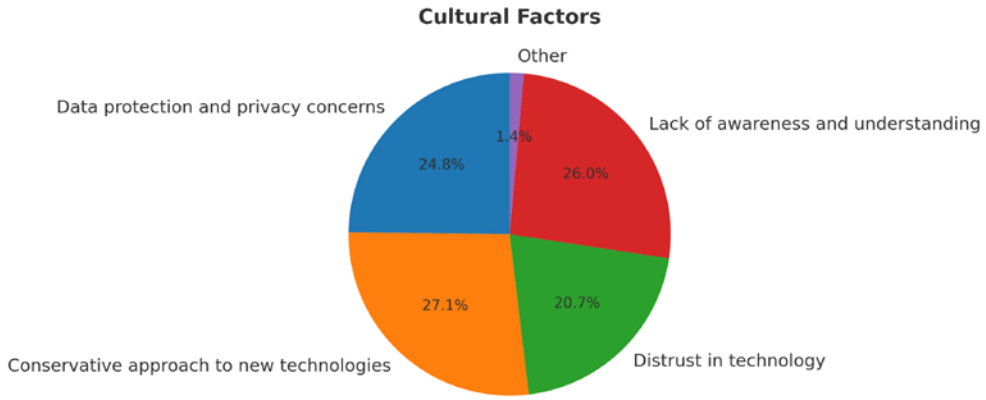


23. In your opinion, what are the main barriers to the application of AI (artificial intelligence) in EHRs (electronic healthcare records) based pharmacovigilance in your country of residence (please select the most significant answers)?

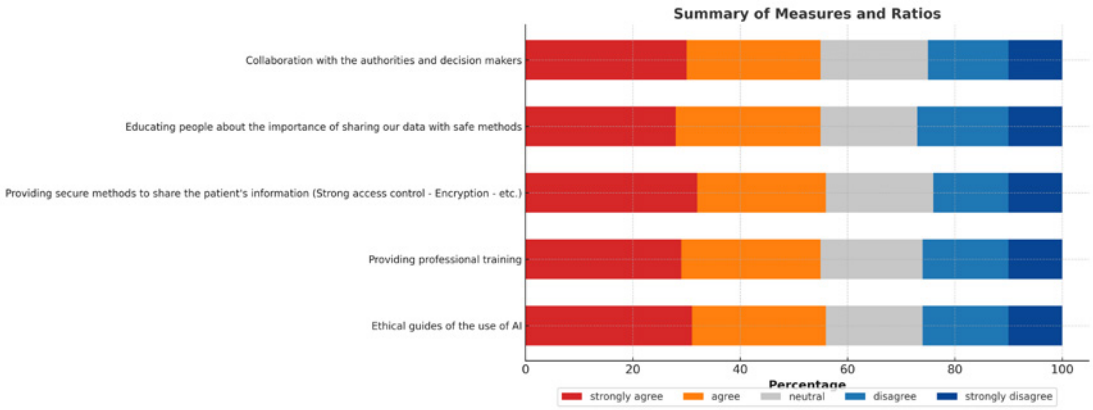


4.4 Challenges and Suggestions

24. Which of these cultural factors are considered barriers to the implementation of AI in EHR-based pharmacovigilance (select the most significant answers)?



25. How will these measures help to increase the implementation of AI and EHR in pharmacovigilance?



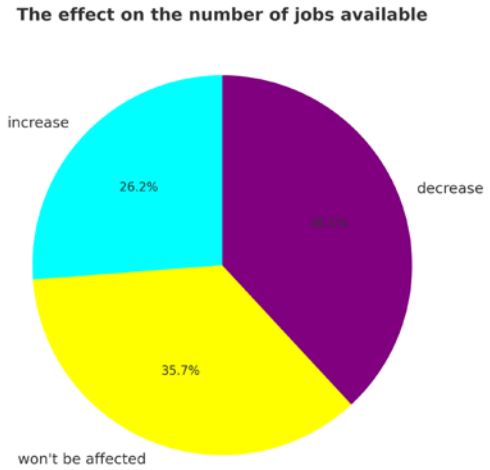
Summary of Measures and Ratios

Measure	Strongly Agree (%)	Agree (%)	Neutral (%)	Disagree (%)	Strongly Disagree (%)
Collaboration with authorities and decision-makers	53.5	35.5	10.5	0.0	0.6
Educating people about sharing data (safe methods)	45.9	37.2	14.5	1.7	0.6
Providing secure methods to share patient information	55.2	33.7	10.5	0.0	0.6
Providing professional training	54.1	32.6	12.8	0.6	0.0
Ethical guides for the use of AI	52.3	30.2	15.7	1.2	0.6

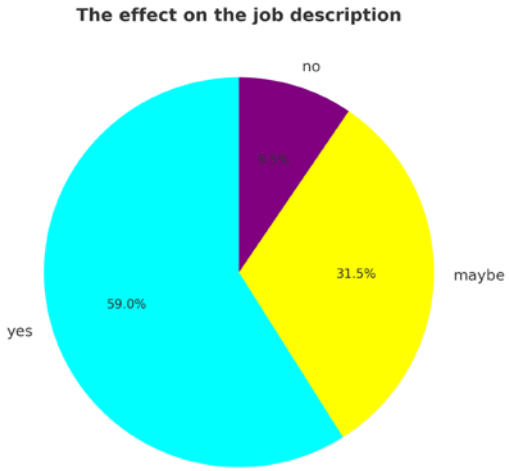
26. Do you have any further suggestions/solutions to enhance the application of AI and EHR in Pharmacovigilance?

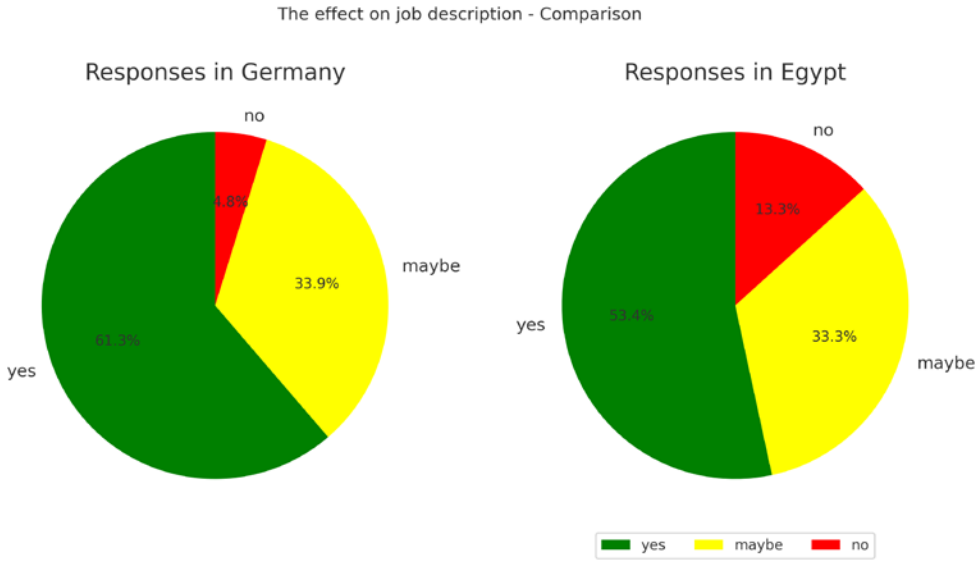
It was an optional question, and we received 62 responses. The significant responses will be added to the Conclusion part.

27. How do you expect the number of jobs available in the PV market to change in the case of the application of EHR and AI in PV?



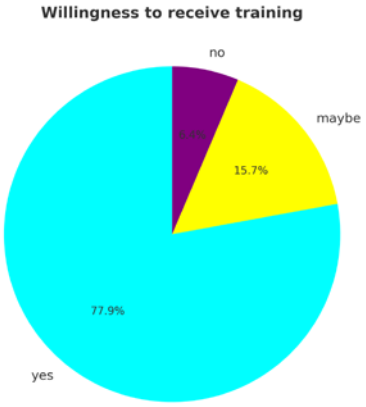
28. Do you expect the role/ job description of a PV specialist to change in the case of the application of AI and EHR in PV in the future?

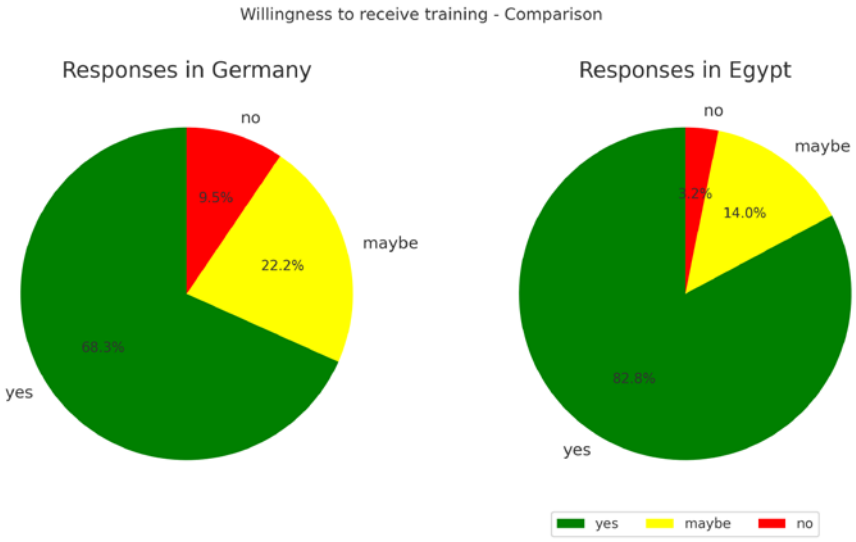




For this question, the two-sample z-test was conducted.
The power of the experiment was 0.164185261766375.
The z-statistic was 0.9726721278116287, while the p-value was 0.33071628210080184
The p-value was approximately 0.331, which was greater than or equal to the prechosen significance level of 0.05. Therefore, we failed to reject the null hypothesis.
There is not sufficient evidence to suggest that there is a significant difference between the proportion of positive responses from Germany (0.613) and Egypt (0.533).

29. Are you interested in receiving training about the use of AI in EHR-based pharmacovigilance?

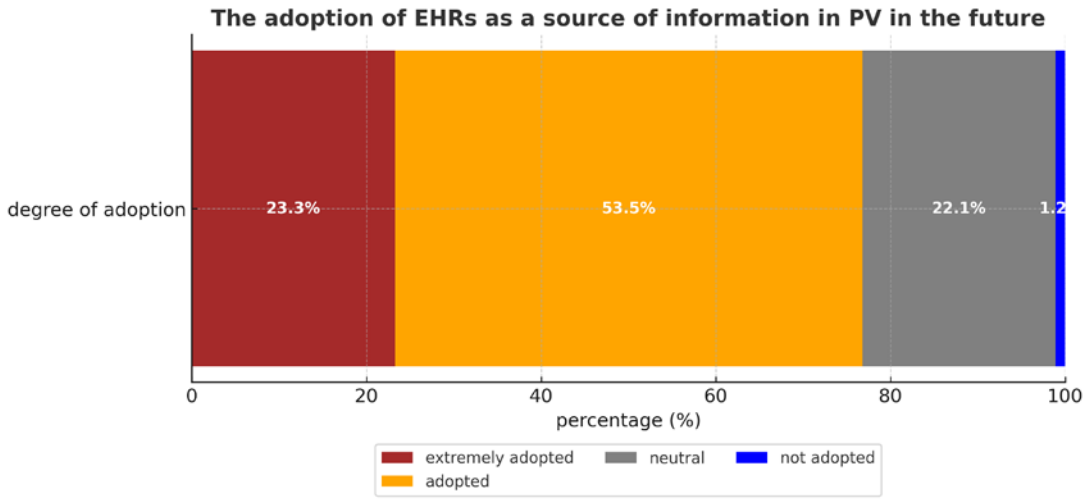




The two-sample z-test was conducted for this question.

The power of the experiment was 0.5533108329740837.
The z-statistic was -2.115184578952405, while the p-value amounted to 0.034414218753534384.
The p-value was approximately 0.034, which was less than the prechosen significance level of 0.05. Therefore, we rejected the null hypothesis. There was sufficient evidence to suggest that there was a significant difference between the proportion of positive responses from Germany (0.683) and Egypt (0.828). The proportion of "yes" responses from Egypt was higher.

30. How do you see the adoption of EHR as a source of information for PV in the next 10 years?



5. Conclusion

5.1 Summary of results

5.1.1 Reality and Application in Germany and Egypt

There are generally positive attitudes toward the adoption of AI and EHRs in the domain of pharmacovigilance **(questions 20 and 29)**. This attitude is not a special phenomenon because of the young age of most participants **(question 2)**.
Most of the participants are PV experts **(question 4)** who have different job descriptions and work in different settings. The result of this question indicates how the use of AI can enrich the PV industry because of the diversity of the tasks and roles in this domain.
Most of the participants are a mix of early-career (1–5 years) and mid-career professionals (6–10 years). The percentages of these two groups were 49.4 % and 26.74 %, respectively (see **question 5**). Most of the participants work in pharmaceutical companies (55.81 %), but there is a diversity in their workplaces as well **(question 6)**. This diversity provides different perspectives and points of view on how AI can be used in PV.

As far as the **8th question** is concerned, EHRs started to be implemented in Germany (still in the initial phase), but not yet in Egypt.
Only 49 % of the participants (in both countries) have heard about the use of EHRs in PV before (question 9).

Regarding the organizations (question 10), most of the organizations in Germany and Egypt are not utilizing EHRs in pharmacovigilance.
When it comes to the resources of information in the domain of PV in both countries, ICSR (individual case safety reports) remain the most important source of data in PV. The literature is still a valid source of information (2nd place), but the local literature review in Egypt is of poor quality. There is a rise in the usage of digital portals and social media, but EHRs are rarely used (question 11).
PV specialists lack knowledge and awareness of the use of AI in PV (question 12). These results indicate the importance of further education and efforts to understand the role of AI in improving PV. In addition, AI is not yet widely applied in the organizations involved in PV activities neither in Germany nor Egypt, and there is no significant difference between the two countries (question 13).

This indicates a generally similar attitude towards the acceptance of AI in PV in both countries.

The most common uses of AI in PV (very few participants answered to this question) are case reporting processing and narrative writing (14 responses). Automated ADR detection (from structured and unstructured data, e.g. EHR, social media, medical literature) and signal detection (by detecting trends and patterns) are some other uses of AI in PV.
AI can be used in the domain of PV to facilitate and accelerate the reporting of side effects on the national pharmacovigilance platforms (e.g. in Germany www.nebenwirkungen.bund.de). AI-supported input assistance could improve the quality of the data by using plausibility checks to directly correct input errors for that reporting, therefore, fewer questions will be required later. Context-dependent additional information can be requested depending on patients' characteristics such as age, gender, or pregnancy, as well as specific symptoms and progression [12].

The most used AI technologies in PV are ML, AI-powered automation, NLP, generative AI, and chatbots, respectively, taking into consideration that very few participants responded to this question (question 15).
There is very little simultaneous use of AI and EHRs in PV organizations in both Germany and Egypt, and there is no significant difference in the adoption between Germany and Egypt in this regard (question 16). There were also very few participants.

In Germany as well as in Egypt, MedDRA is the most known and used coding language in PV **(question 17)**. The ICD10 comes in second place. Although few PV experts are familiar with other concepts of medical terminology and the standards of data share and interoperability, there is still a lack of knowledge of these domains among PV experts. The WHO Drug Dictionary is also used.

Regarding the restrictive following of the data protection regulations in Egypt and Germany **(question 18)**, the results show that PV organizations in both countries do follow the rules of data protection. However, a comparative analysis of the survey's data shows that most PV organizations in Egypt do not strictly comply with data protection and privacy regulations in comparison to Germany. There is far greater adherence in the PV organizations in Germany. These findings highlight potential disparities in the implementation of data protection measures between the two countries, which could impact on the handling and processing of sensitive healthcare data, including data used in PV practices involving AI technologies. So, there is a serious problem in the data protection laws in Egypt which demands interventions of the decision-makers.
Most of the participants received training about the software used in pharmacovigilance in our workplace, but a few of them received formal training in digital transformation. There is no significant difference between the two countries in this regard **(question 19)**.

The survey results **(question 20)** show positive attitudes from the participants toward the adoption of **EHRs and AI** in the field of pharmacovigilance. This reflects the specialists' awareness of the potential positive impact of these technologies on the field of PV. These positive attitudes and perceptions should encourage the decision-makers to introduce these technologies and provide financial support for the companies and the required training for pharmacovigilance experts.

5.2 Potential Opportunities in Germany and Egypt

Based on the results of the survey, there are a lot of expected benefits of the simultaneous use of AI and EHRs in PV **(question 21)**. The implementation of these technologies will provide us with more data to be analyzed and will help to solve the eternal problem of **underreporting and the poor quality of reports**. There are other expected benefits like the facilitation of narrative writing, a faster signal detection process, automated case triage, more accuracy, reduced human error, improved ADRs, increased compliance with regulations, faster data analysis, and increased patient safety.
The use of AI in EHR-based pharmacovigilance will have a positive impact on the health status of the country **(question 22)**. Some of these benefits are increased safety and efficacy of drug therapy,

enhanced PH (public health), reduction of the hospitalization rate due to ADRs, improved patients' quality of life, contribution to the cost-effectiveness of health services, and enhanced research for rare diseases.

5.3 Challenges and Barriers in Germany and Egypt

There are a lot of barriers to the implementation of AI in EHR-based PV **(question 23)**. The barriers are the following: cultural barriers, the fear of job displacement, resistance from the staff to change, data privacy and patient data protection (refusal of the use of clinical data), technical barriers (data quality issues – unstructured data in EHRs – lack of AI expertise – reliability of AI algorithms – data exchange problems), political barriers, lack of personnel training and lack of resources. The most significant barriers are the technical barriers, then secondly the lack of resources and data privacy and protection, then the fear of job displacement and the resistance to change from staff.

One of the major technical problems is the low quality of data and unstructured data. This problem can be approached through the standardization of medical terminology and the mapping between different medical languages, in addition to the use of standards of interoperability between different systems to share data [30, 31]. Another approach to deal with this challenge is to develop and to update the medical languages (i.e. MedDRA). Also, sequential labeling of the data in EHRs is crucial, in addition to the definition of the metadata and the use of common data models (CDM).
To enhance the performance of ML, high-quality data should be used. In addition, the effectiveness of existing AI algorithms necessitates a "human-in-the-loop".

Regulatory barriers are crucial barriers as well; to overcome these barriers, we need to have new **PV regulations** that permit the use of AI rather than the traditional regulations that impede its implementation. In addition, **regulations regarding AI use** and regulations regarding the secondary use of data are required. Another suggested solution is to adapt the digitalization plan of the healthcare systems (e.g. Digital Health Act in Germany). Financial support is also crucial to solving the problems of lack of training and the fear of job displacement.

The conservative approach to new technologies (140 responses) is the most significant cultural barrier toward the implementation of AI in EHR-based PW. The lack of awareness and understanding (e.g. lack of patient reporting) comes in second place (134 responses). Pharmacovigilance is only included in the syllabus of pharmaceutical studies, while other medical curricula do not convey any knowledge of drug safety in general, especially in pharmacovigilance [23]. Then comes data protection and privacy concerns (128 responses), and finally the distrust in technology (107 responses; **question 24**).

There are a lot of measures that may increase the implementation of AI and EHRs in PV **(questions 25 and 26)**: One of these measures is a collaboration with the authorities and decision-makers, proposing the road map, and having a long-supported pilot phase. The roadmap for the implementation of the digital health law in Germany can be a role model [32].

Educating people about the importance of sharing our data (with safe methods) is crucial to overcoming misunderstandings and fears. One of the ways to address people's fear is to provide ethical guidelines for the use of AI, while at the same time providing secure methods to share the patient's information, e. g. strong access control – encryption – pseudonymization – anonymization or blockchains [33 ,34]. Providing professional training for PV experts is crucial.

Regarding data privacy and protection, we need to apply the rules correctly, and at the same time, we need to ensure feasible access to health data for the sake of usage in PV and other research purposes. EHDS is a promising project to solve this dilemma [35].

Initiatives that encourage stakeholder collaboration for diverse datasets are crucial. Patient engagement can be increased by leveraging AI-driven tools to engage patients in PV efforts, such as mobile apps

for reporting adverse events and monitoring medication adherence. Raising the awareness of the stakeholders and the decision-makers is also crucial.

There was no absolute agreement regarding the number of jobs available in case AI and EHRs are applied in the domain of PV **(question 27)**. Some see that AI will replace conventional human tasks and others see that the implementation of these cutting-edge technologies will create new opportunities. Despite those fears, there are still positive attitudes toward AI and EHRs. The rise of awareness and the motivation to develop will help to reduce such fears.
A majority of the participants expect a change in the role description of PV specialists, and these expectations are similar, with no significant difference between Egypt and Germany **(question 28)**.
There is a great interest in receiving training about the use of AI in EHR-based PV in both countries, and this interest is greater in Egypt than in Germany. Germany may be more conservative toward new technologies, in addition to the competitive nature of the market in Egypt **(question 29)**.
Finally, the participants show a positive attitude toward the adoption of EHRs as the source of information for PV in the next 10 years. This indicates the awareness of the positive impact of EHRs as a source of information on PV **(question 30)**.

5.4 Recommendations

Egypt can profit from the road map of the German government regarding the digitalization of the healthcare system, minimally starting with a good documentation system. In addition, Egypt should issue a comprehensive law on data protection following the steps of the European Union. These two pillars, namely, the digital healthcare system and data protection laws, are indispensable for taking further steps toward the secondary use of data for PV and research purposes.
Digital health specialists also need to raise awareness in the public, towards healthcare professionals and the regulatory authority about the importance of digital tools in the domain of PV.
In addition, the regulatory bodies in both countries need to profit from the positive attitude toward these technologies (AI, EHRs) and accelerate the process of integration of these technologies in healthcare generally and in PV specifically. This includes offering training and providing financial aid to the PV companies.

5.5 Limitations

- One of the limitations of these studies is the sample size. Although the sample size is quite acceptable, if the researcher gets more respondents he will have a sample size calculation with a better confidence interval and margin of error.
- Some of the key stakeholders in the domain of PV refused to participate in the survey for different reasons.
- The difficulty in accessing real-world data in EHRs to test the ability of AI to detect ADRs in EHRs. This difficulty could be attributed to the regulatory and privacy barriers. Both research datasets (which are available for research purposes) and synthetic data were insufficient because they do not fully present the reality.
- In addition, the technical barriers to the possibility of implementation of AI and EHRs would be better discussed with IT specialists with special interviews.

5.6 Suggestions for future research

These are some of the suggested areas for future research:

- The use of AI in community pharmacies: How can AI enhance the role of the pharmacist regarding public health?
- The use of AI in clinical research (clinical trials)
- The use of AI in drug design
- Interoperability standards/mapping for EHR-based PV to synchronize the PV activities with EHRs.
- Security methods for EHR-based pharmacovigilance
- Facilitation of the restricted rules for reporting and the issuing of new legislation for the different types of PV.
- How can the use of EHRS enhance research on rare diseases?
- The use of AI in social media-based PV.
- Systematic reviews of the current state of PV in Egypt and Germany.
- The barriers toward the digitalization of healthcare systems in Germany and Egypt.

Acknowledgement

Special thanks to the Eng. Rafiq Mazen for his guidance and help through the process of data analysis.

Conflicts of Interest Statement

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] Shin 6H, Lee S. An OMOP-CDM based pharmacovigilance data-processing pipeline (PDP) providing active surveillance for ADR signal detection from real-world data sources. *BMC Medical Informatics and Decision Making.* 2021;21(1):159. https://doi.org/10.1186/S12911-021-01520-Y.

[2] Centers for Disease Control and Prevention (CDC). Adverse Drug Events in Adults | Medication Safety Program | CDC. CDC. 2023. https://www.cdc.gov/medication-safety/data-research/facts-stats/index.html. Accessed April 6, 2024.

[3] European Commission. Pharmacovigilance – European Commission. European Union. 2024. https://health.ec.europa.eu/medicinal-products/pharmacovigilance_en. Accessed April 5, 2024.

[4] Cliniskills. Data Sources in Pharmacovigilance – ClinSkill. Cliniskills Digital Learning Academy. 2023. https://www.clinskill.com/docs/data-sources-in-pharmacovigilance/. Accessed April 6, 2024.

[5] Trifirò G, Sultana J, Bate A. From Big Data to Smart Data for Pharmacovigilance: The Role of Healthcare Databases and Other Emerging Sources. *Drug Safety.* 2018;41(2):143–149. https://doi.org/10.1007/S40264-017-0592-4/METRICS.

[6] Analytics P. Sources of data – Pharmacovigilance Analytics. PHARMACOVIGILANCE ANALYTICS. 2018. https://www.pharmacovigilanceanalytics.com/pharmacovigilance-analytics/sources-of-data/. Accessed April 6, 2024.

[7] European Medicines Agency (EMA). Launch of new HMA-EMA catalogues of real-world data sources and studies | European Medicines Agency. EMA. 2024. https://www.ema.europa.eu/en/news/launch-new-hma-ema-catalogues-real-world-data-sources-studies. Accessed April 6, 2024.

[8] Yu Y, Ruddy KJ, Hong N, et al. ADEpedia-on-OHDSI: A next generation pharmacovigilance signal detection platform using the OHDSI common data model. *Journal of Biomedical Informatics*. 2019;91:103119. https://doi.org/10.1016/J.JBI.2019.103119.

[9] Salgado CM, Azevedo C, Proença H, Vieira SM. Noise versus outliers. InMIT Critical Data. *Secondary Analysis of Electronic Health Records.* Published online 2016:163-183.

[10] Sauzet O, Cornelius V. Generalised weibull model-based approaches to detect non-constant hazard to signal adverse drug reactions in longitudinal data. *Frontiers in Pharmacology.* 2022;13:13:889088. https://doi.org/10.3389/FPHAR.2022.889088.

[11] Davis SE, Zabotka L, Desai RJ, et al. Use of Electronic Health Record Data for Drug Safety Signal Identification: A Scoping Review. *Drug Safety.* 2023;46(8):725–742. https://doi.org/10.1007/S40264-023-01325-0.

[12] Schuster N. Datenberge mit KI bewältigen. *Pharmazeutische Zeitung.* 2024;(6):18.

[13] Son N, Kim B, Chung S, Han S. Korean pharmacovigilance system based on EHR-CDM. *Studies in Health Technology and Informatics.* 2019;264:1592–1593. https://doi.org/10.3233/SHTI190550.

[14] Bonomi S. The electronic health record: A comparison of some European countries. *Lecture Notes in Information Systems and Organisation.* 2016;15:33–50. https://doi.org/10.1007/978-3-319-28907-6_3/COVER.

[15] Organisation for Economic Co-operation and Development (OECD). Digital health | Health at a Glance 2021 : OECD Indicators | OECD iLibrary. OECD iLibrary. 2019. https://www.oecd-ilibrary.org/sites/08cffda7-en/index.html?itemId=/content/component/08cffda7-en. Accessed March 23, 2024.

[16] Yu Y, Ruddy KJ, Wen A, et al. Integrating Electronic Health Record Data into the ADEpedia-on-OHDSI Platform for Improved Signal Detection: A Case Study of Immune-related Adverse Events. 2020;2020:710–719.

[17] Food and Drug Administration (FDA). Questions and Answers on FDA's Adverse Event Reporting System (FAERS) | FDA. FDA. 2019. https://www.fda.gov/drugs/surveillance/questions-and-answers-fdas-adverse-event-reporting-system-faers. Accessed December 3, 2023.

[18] Food and Drug Administration (FDA). FDA's Sentinel Initiative | FDA. FDA. 2023. https://www.fda.gov/safety/fdas-sentinel-initiative. Accessed December 3, 2023.

[19] Douros A, Schaefer C, Kreutz R, Garbe E. Pharmakovigilanz in Deutschland: Es wird langsam Zeit. *Internist.* 2016;57(6):616–623. https://doi.org/10.1007/S00108-016-0068-0.

[20] Gahr M, Eller J, Connemann BJ, Schönfeldt-Lecuona C. Underreporting of adverse drug reactions: Results from a survey among physicians. *European Psychiatry.* 2017;41(S1):S369. https://doi.org/10.1016/J.EURPSY.2017.02.377.

[21] ELkhwsky FS, El Sayed I, Yassine OGM, Abdelmonem S, Salama MM. Healthcare professionals' knowledge and practice of and attitudes towards pharmacovigilance in Alexandria, Egypt: a cross-sectional survey. *Drugs and Therapy Perspectives.* 2021;37(3):124–136. https://doi.org/10.1007/S40267-020-00798-8.

[22] Aziz AA, Rogers S, Hassanien O, Shalaby L, Nagy M. Knowledge, attitudes and practice regarding pharmacovigilance and adverse drug reaction reporting among physicians and pharmacists in Egypt: a step toward personalized medicine implementation. 2022;19(6):495–507. https://doi.org/10.2217/PME-2022-0030.

[23] Kaeding M, Schmälter J, Klika C. *Pharmacovigilance in the European Union. Practical Implementation across Member States*; 2017. https://doi.org/10.1007/978-3-658-17276-3.

[24] Good Calculators. Sample Size Calculator | Good Calculators. Good Calculators. 2024. Accessed May 16, 2024. https://goodcalculators.com/sample-size-calculator/.

[25] Shah C. *A Hands-on Introduction to Data Science.* Cambridge University Press; 2020.

[26] Collingridge D. Validating a Questionnaire — Methodspace. Methodspace. 2022. https://www.methodspace.com/blog/validating-a-questionnaire. Accessed November 30, 2023.

[27] Bundesgesundheitsministerium (BGM). Driving the digital transformation of Germany's healthcare system for the good of patients. BGM. 2020. https://www.bundesgesundheitsministerium.de/en/digital-healthcare-act. Accessed April 23, 2024.

[28] ALL About Pharma Covigilance. Understanding E2B | Pharmacovigilance. allaboutpharmacovigilance. 2021. https://allaboutpharmacovigilance.org/understanding-e2b/. Accessed April 1, 2024.

[29] CIOMS. Pharmacovigilance – CIOMS. CIOMS. 2017. https://cioms.ch/pharmacovigilance/. Accessed April 1, 2024.

[30] Jagannatha AN, Yu H. Bidirectional RNN for medical event detection in electronic health records. *2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL HLT 2016 – Proceedings of the Conference.* Published online 2016:473–482. https://doi.org/10.18653/v1/n16-1056.

[31] Lee S, Han J, Park RW, et al. Development of a Controlled Vocabulary-Based Adverse Drug Reaction Signal Dictionary for Multicenter Electronic Health Record-Based Pharmacovigilance. *Drug Safety.* 2019;42(5):657–670. https:// doi.org/10.1007/S40264-018-0767-7.

[32] Bundesgesundheitsministerium (BGM). Digital together. Germany's Digitalisation Strategy for Health and Care. *Federal Ministry of Health.* 2022;2022(4):12-13.

[33] Mishra R, Ramesh D, Edla DR, Qi L. DS-Chain: A secure and auditable multi-cloud assisted EHR storage model on efficient deletable blockchain. *Journal of Industrial Information Integration.* 2022;26. https://doi.org/10.1016/j.jii.2021.100315.

[34] Hasan QH, Yassin AA, ATA O. Electronic Health Records System Using Blockchain Technology. *Webology*. 2021;18(Special Issue):580–593. https://doi.org/10.14704/WEB/V18SI05/WEB18248.

[35] Evans J. EU stärkt EMA und Datenaustausch. *Pharmazeutische Zeitung*. 2023;(51–52):16.