

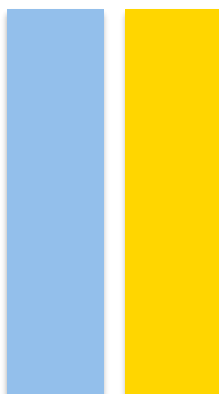
**MESTRADO**  
INFORMÁTICA MÉDICA

# **Evaluation of a pharmacovigilance causality assessment tool**

João Filipe Coutinho de Almeida

**M**

2019





12ª Ed

MIM

# Evaluation of a pharmacovigilance causality assessment tool

MESTRADO EM  
INFORMÁTICA MÉDICA  
2º CICLO DE ESTUDOS

João Filipe Coutinho de Almeida

SET | 2019



12ª Ed

MIM

# Evaluation of a pharmacovigilance causality assessment tool

MESTRADO EM  
**INFORMÁTICA MÉDICA**  
2º CICLO DE ESTUDOS

João Filipe Coutinho de Almeida

ORIENTADORES:

Pedro Pereira Rodrigues,  
Inês Ribeiro-Vaz, Ricardo João Cruz Correia

SET | 2019

To my parents for always  
supporting me through my education  
and to Lia for the unlimited patience.

## Resumo

As reações adversas ao medicamento (RAM) são uma grande preocupação para todos os profissionais de saúde, entidades governamentais e para a sociedade em geral. Para conseguir colmatar este problema, as entidades acabam por depender muito de notificações voluntárias da população. Deste modo, é estrutural promover a importância da notificação na população e encontrar métodos de melhorar o processo e sistematizar métodos.

Foi neste enquadramento que se começou a desenvolver uma ferramenta de suporte à avaliação de causalidade na Unidade de Farmacovigilância do Porto (UFP). Foi pelo entendimento que os centros de farmacovigilância portugueses e mundiais dependem de imenso tempo a recolher informações por cada notificação de suspeita de reação que recebem e devido ao facto de ser algo já muito sistematizado, poderia ser automatizado. O sistema coleciona informações de 4 *websites* muito usados na compilação de informação – *vigiaccess*, EMA, MHRA e *medscape*. A ferramenta foi colocada num ambiente controlado de forma a obter dados e *feedback* dos utilizadores finais.

Com isto, propõe-se então nesta tese avaliar o estado atual da ferramenta, perceber que melhorias são fundamentais para a sua plena implementação. É igualmente um objetivo recolher *feedback* de profissionais de saúde que são incluídos no ciclo do medicamento de forma a entender as razões de números baixo de notificação e como/se a tecnologia poderá ter um impacto positivo nesta questão. A par disto, perceber o estado da arte da informática na farmacovigilância, de modo a suportar o desenvolvimento da aplicação, obter melhores práticas e diminuir esforços de desenvolvimento.

## **Abstract**

Adverse drug reactions (ADR) are a major concern for healthcare professionals, government, and society. Portuguese pharmacovigilance centres spend immense time collecting information every time a notification of suspected ADR is made. Our research team developed an early concept of a tool for helping pharmacovigilance teams improve the causality assessment process, in order to increase the speed overall, better and faster feedback for reporters and overall better encouragement for notifications.

The tool is currently based on a system that is able to collect ADR information from several sources and store it in a structured format, in order to decrease time spent, by the Pharmacovigilance team, looking up for information, while providing better and faster information.

In this work, we propose to assess the current state of the tool, gathering feedback from Key opinion leaders of the healthcare and pharmacovigilance background. This will permit a better bird's eye view of the tool, current strengths and weaknesses and suggest a road map for its further development.

# Table of Contents

Acknowledgements .....	<b>Error! Bookmark not defined.</b>
Resumo .....	V
Abstract .....	VI
Table of Contents .....	VII
Table Index .....	IX
Figure .....	X
Acronym Index.....	XI
1. Introduction.....	1
1.1. Pharmacovigilance in Portugal .....	2
1.2. The need for reporters.....	5
1.3. The causality assessment support tool.....	6
1.4. Pharmacovigilance Informatics: state of the art .....	8
1.5. Objectives .....	11
2. Assessment of biomedical systems .....	12
2.1. What is evaluation? .....	13
2.2. Evaluation Methods.....	16
2.2.1. Quantitative studies .....	17
2.2.2. Qualitative studies.....	20
3. Design & Methods .....	24
3.1. What capabilities should a causality assessment support system have in order to improve report numbers.....	24
3.1.1. Research methods .....	25
3.2. Is the current system able to act upon expected? .....	28
3.2.1. Methods.....	29
4. Results.....	30
4.1. Interviews results.....	30
4.2. Feature evaluation .....	30
5. Analysis and Discussion .....	35
5.1. Future developments and insights .....	35
5.2. Interviews.....	37
5.2.1. Notification of ADR.....	37
5.2.2. Information Systems for helping ADR notification.....	38
5.2.3. Roadmap.....	39
5.3. Technical capabilities .....	39

6. Conclusions.....	41
7. References.....	42
ANNEXES.....	49
ANNEX A – Images of the causality assessment tool .....	49
ANNEX B – Images of SINAVE .....	51
ANNEX C – Pharmaceutical Forms without translation .....	52

## Table Index

Table 1 Different Study Types .....	14
Table 2 Steps for creating successful interviews.....	27
Table 3 Performance Indicators of the system .....	30
Table 4 Results of non-successful searches in the system .....	30

# Figure Index

Figure 1 ADR reporting workflow for Porto Pharmacovigilance Unit .....	4
Figure 2 Quantity of spontaneous reports in Portugal per origin, adapted from Infarmed I.P. (30) .....	5
Figure 3 Potential of different data to be linked together .....	9
Figure 4 Possibilities of task automation in pharmacovigilance from Friedman et al. (41) .....	10
Figure 5 Potential Uses of blockchain in healthcare .....	11
Figure 6 Objectivist study model .....	17
Figure 7 Relation between measurement and demonstration studies (66).....	19
Figure 8 Subjectivist Study model .....	22
Figure 9 Time of search for different sites and requests .....	33
Figure 10 Loading times for reaction's groups .....	33

## Acronym Index

<b>ADE</b>	Adverse Drug Event
<b>ADR</b>	Adverse Drug Reaction
<b>API</b>	Application Programming Interface
<b>CIOMS</b>	Council for International Organizations of Medical Sciences
<b>EMA</b>	European Medicines Agency
<b>ER</b>	Emergency Room
<b>EU</b>	European Union
<b>FDA</b>	Food and Drug Administration
<b>FHIR</b>	Fast Healthcare Interoperability Resources
<b>HTML</b>	Hyper Text Markdown Language
<b>IDMP</b>	Identification of Medicinal Products
<b>INFARMED, I.P.</b>	National Authority of Medicines and Health Products
<b>JSON</b>	JavaScript Object Notation
<b>KOL</b>	Key Opinion Leaders
<b>KPI</b>	Key Performance Indicator
<b>MA</b>	Marketing Authorization
<b>MAH</b>	Marketing Authorization Holder
<b>EEA</b>	European Economic Area
<b>MHRA</b>	Medicines and Healthcare Products Regulatory Agency
<b>MVP</b>	Minimum Viable Product
<b>NoSQL</b>	Not Only Structured Query Language
<b>NPS</b>	National Pharmacovigilance System
<b>REST</b>	Representational State Transfer
<b>UK</b>	United Kingdom
<b>UMC</b>	Uppsala Monitoring Centre
<b>USA</b>	United States of America
<b>WHO</b>	World Health Organization
<b>WWID</b>	Worldwide ID

# 1. Introduction

Adverse drug reactions (ADR) are defined by the WHO as a noxious or unintended response to a medicinal product (1). A related term, adverse drug event (ADE), is defined as any event that occurs during treatment, and is not necessarily caused by the drug itself (2). This definition includes harm caused by the drug (ADRs and overdoses) and harm from the use of the drug (including dose reductions and discontinuations of drug therapy) (3,4). ADE are the most common cause of iatrogenic harm in healthcare (5) and virtually impossible to assess until the drug commercialization.

ADR are classified into several types according to a few criteria. Type A are augmented, meaning that is an expected behaviour for the drug, providing the systemic change that is supposed to provoke, but the effects are greater than desired on a certain patient. This issue is resolved by adjusting the dose. Type B reactions are not pharmacologically expected and are individually-dependent. This issue is corrected by administration suspension. Reactions of type C are related to a cumulative effect of the drug in the system and are only visible after a prolonged intake of a drug. Type D are very connected to a drug that was manipulated into a suspension and mostly appear after a very long period of time, meanwhile Type E are similar to D, but occur after a short time of the suspension formulation. Finally Type F occur by a lack of efficacy of a drug (6).

Reports available in the last years shed light into major issues regarding these reactions, referring that in average, 6.7% of hospitalizations worldwide are due to some kind of ADR (7). More recently, these values have been detailed and indicated that serious ADR can vary from as low as 0.2% to as high as 64%, depending on country and institution, reflecting a possible environment – linked issue and that most of the ADR were type A, therefore potentially avoidable (8). As for the national reality, reports say that over 18 thousands ADR occur every year in the Portuguese hospitals, increasing yearly over the last 20 years (9).

ADR not only are responsible for major healthcare impacts, but also for representing a great economic burden. This is mainly due to the impact of undocumented ADR in the safety of drugs commercially available and the financial burden that these events imply. Some examples of some measured impacts are:

- the responsibility for one-third of all adverse events during hospital stays, affecting 2 million stays annually (10) in the United States of America (USA);
- related to over 3.5 million physician office visits and 1 million Emergency Room (ER) visits (11) in the USA;
- responsibility for 3.5 billion dollars in health care costs (12) in the USA .
- responsibility for 2.5 billion dollars cost in health care in the United Kingdom (UK) (13);

Specifically, for European countries, findings suggest similar values, such as a conclusion provided by the European Commission that estimates that approximately 5% of all hospital admissions are caused by ADR, that 5% of hospitalized patients will experience an ADR during their hospital stay and that ADR cause 197 000 deaths in the European Union (EU) (14). These findings were decisive for the reform of the European

regulatory system of the EU in 2012 and further studies have sustained these numbers (15).

In Portugal, for a drug to get into the Portuguese market, a Marketing Authorization (MA) must be obtained. For this, several additional studies and trials are conducted in order to provide the most accurate and supportive evidence regarding safety and efficiency. Nevertheless, there are some ADR that may not be caught during these trials and that is why ADR are assumed as an inherent risk of drug therapies since it is only possible to fully verify the most reactions associated with a drug after a few years in the market. This issue is derived from different reasons, like results of testing drugs on animals do not always correlate to those obtained when testing the same drug on humans. Secondly, the patient population recruited during clinical trials is small and biased and hence the data are not statistically robust. In fact, clinical trials fail to identify rare and serious side-effects, due to relatively small study duration (16), so there are some ADR that simply do not occur during trials. They can be very rare or take a very long time until they are detected. For instance, are currently described ADR that have a frequency of 1 in 10 thousand patients or less (6). These types of ADR are only possible to monitor and describe after the drug is available to the general public.

It is with these issues in mind that pharmacovigilance was created. It started its path after the thalidomide tragedy of 1961, where a improperly tested drug which sold as a nonaddictive, nonbarbiturate sedative for pregnant women, resulted into over 10 000 of birth defects and increased miscarriage rates (17,18). Since then, it has evolved as a result of collaborative efforts by physicians, pharmacists, healthcare providers, patients, health and medicine authorities, academia, industry, the World Health Organization (WHO), the Council for International Organizations of Medical Sciences (CIOMS), and the International Conference on Harmonization (19). It is currently defined by the WHO as “the science and activities relating to the detection, assessment, understanding, and prevention of adverse effects or any other drug-related problems” (20). This science is a major activity for patient and drug safety, and is mainly carried out for pharmaceutical products after they are introduced to the market. Pharmacovigilance presently heavily relies on the collection and analysis of spontaneous reports of adverse events obtained from drug manufacturers, healthcare professionals, and consumers. With the exclusion of manufacturers, reporting into this system is voluntary, which is why the process is often characterized as passive (21). Spontaneous reporting systems have proven vital to post-marketing surveillance and notification increase, and are effective at detecting many types of ADRs, especially rare ones. However, the significant delays in detecting other types of ADRs, and the realization that a substantial number of ADRs remain unreported, have led to the search of complements for the current pharmacovigilance systems (22).

### 1.1. Pharmacovigilance in Portugal

In Portugal, the National Pharmacovigilance System (NPS) was established by law in 1992, and headed by the National Authority of Medicines and Health Products (INFARMED, I.P.). Although the NPS was established in a centralized manner, it soon became apparent that geographic decentralization would be advantageous and the result was an arrangement of four regional pharmacovigilance units (North, Centre, Lisbon and South) for covering the entire mainland (23,24). Nowadays there are even

more units, summing to a total of 9 regional offices in Portugal – Guimarães, Porto, Coimbra, Beira Interior, Setúbal e Santarém, Alentejo, Algarve, Açores and Madeira.

These centres are responsible for receiving and processing the suspected ADR spontaneous reports. These can be made through the national medicine authority or through the regional pharmacovigilance centres, by paper, telephone or online. ADR reports are received by the regional centres, sent to the national medicine authority, which sends the information to the European Medicine Agency (EMA), World Health Organization's (WHO) Uppsala Monitoring Centre (UMC) and the Marketing Authorization holder (MAH).

Using the Porto pharmacovigilance unit as reference, when an ADR report arrives at the centre, the pharmacovigilance staff trigger a workflow in order to provide an adequate response. The report is reviewed as some of the information received is validated, such as the type of adverse reactions reported and the degree of seriousness attributed by the reporter. If more information is required, the reporter is contacted and any doubts are cleared. The ADR report is then transcribed into the INFARMED ADR Database. The insert into the national database, gives the event a worldwide ID (WWID), identifying this report among all the reports around the world. Afterwards, a manual search in several websites is done, aiming to compile information about the case at hand. It is checked if the suspected ADR has already been described and its frequency (if found). The main websites used are:

- Medscape.com, a private owned site, that provides Drug Monographs, with information relative to adverse effects;
- MHRA.UK, which is the official website for the Medicines & Healthcare products Regulatory Agency – English Medicine Authority. This website is supported by the Yellow Card Scheme, the English ADR reporting system and contains an overview of all UK suspected ADR reports;
- ADRreports.eu, EMA's website for looking suspected adverse drug reactions for authorized medicines in the European Economic Area (EEA).
- VigiLyze and VigiAccess, which are sites powered by the WHO and provide information about suspected ADR from all over the world. VigiLyze are for properly authenticated pharmacovigilance officers and VigiAccess is open to the population.

After the information is compiled, a report is filled and delivered to the physician working at the centre for causality assessment. This process is done through Global Introspection, in which an expert (or a group of experts) expresses judgement about possible drug causation, considering all available data in the ADR report. The decision is based on the expert knowledge and experience, and uses no standardized tools. The result of the causality assessment is categorized into levels of certainty defined by WHO which are as follow.

**Certain:**

- Event or laboratory test abnormality, with plausible time relationship to drug intake;
- Cannot be explained by disease or other drugs;
- Response to withdrawal plausible (pharmacologically, pathologically);
- Event definitive pharmacologically or phenomenologically (i.e. an objective and specific medical disorder or a recognized pharmacological phenomenon);
- Rechallenge satisfactory, if necessary;

**Probable / Likely:**

- Event or laboratory test abnormality, with reasonable time relationship to drug intake;
- Unlikely to be attributed to disease or other drugs;
- Response to withdrawal clinically reasonable;
- Rechallenge not required;

**Possible:**

- Event or laboratory test abnormality, with reasonable time relationship to drug intake;
- Could also be explained by disease or other drugs;
- Information on drug withdrawal may be lacking or unclear;

**Unlikely:**

- Event or laboratory test abnormality, with a time to drug intake that makes a relationship improbable (but not impossible);
- Disease or other drugs provide plausible explanations;

**Conditional / Unclassified:**

- Event or laboratory test abnormality;
- More data for proper assessment needed;
- Additional data under examination

**Unassessable / Unclassifiable**

- Report suggesting an adverse reaction;
- Cannot be judged because information is insufficient or contradictory;
- Data cannot be supplemented or verified;

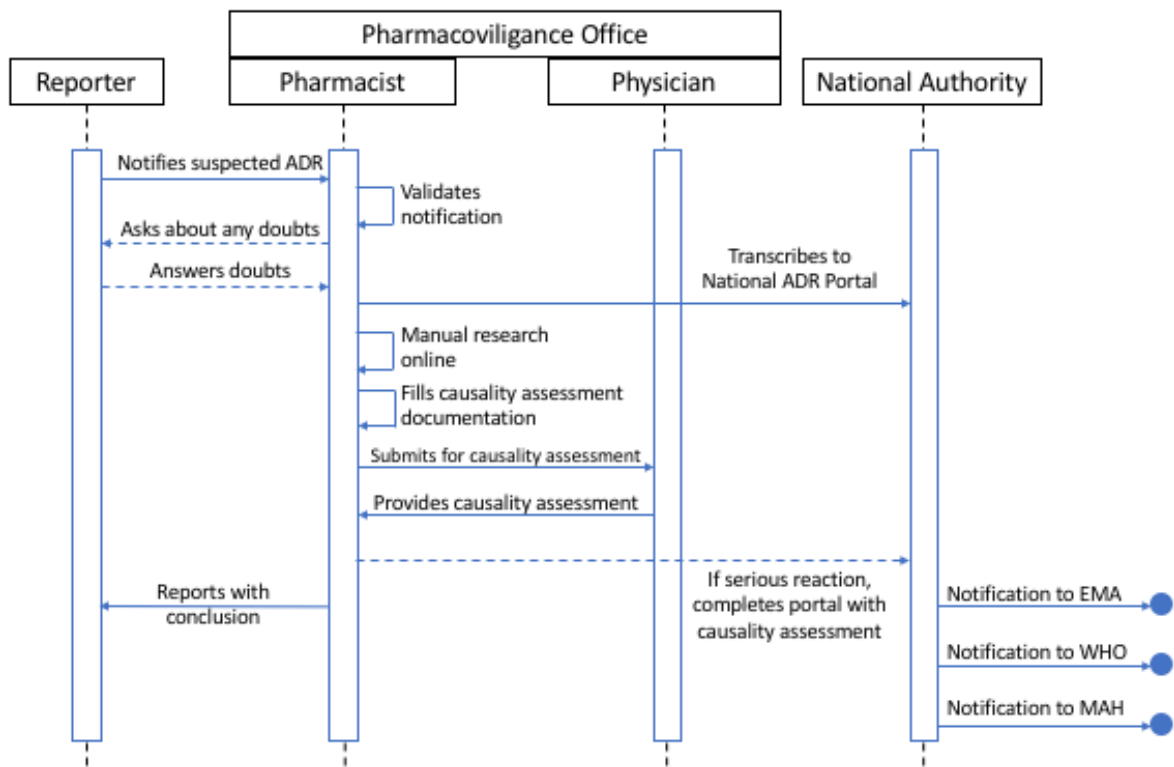


Figure 1 ADR reporting workflow for Porto Pharmacovigilance Unit

For reports related to serious events, the INFARMED ADR database is accessed again and the result of causality assessment is added to the event. Finally, the reporter is

contacted again with a small summary of the event, and the result of the causality assessment given to the case.

### 1.2. The need for reporters

Like stated above, most pharmacovigilance systems are based on spontaneous reporting of ADR (25) and the Portuguese one is no exception. Even though some criticism exist regarding the pertinence of this method and its known limitations, spontaneous ADR reporting remains one of the greatest tools for creating a drug's safety profile (26,27). Regarding the limitations, spontaneous ADR reporting suffers from incomplete data and, in many cases, inadequate data quality. Besides, it depends on the motivation of healthcare professionals and consumers to report, which is not always easy to promote and incentivize. The consequence of this is underreporting, which is a huge problem in all developed countries. It is estimated that only about 10% of the ADR that occur are effectively reported to the regulatory authorities (25,27). Moreover, even though the major interested party are the patients, reports are mainly done by healthcare professionals. In Portugal, according to the latest data from INFARMED, I.P., the main reporters are the pharmacovigilance officers in the pharmaceutical industry (because they are legally bound to do so), followed by physicians and with pharmacists with the third place as can be seen in the picture below (28). One value of notice is the dramatic increase of notifications done by the industry in 2018. This is explained by paradigm shift induced by the notification regulation in November 2017, which created a mandatory status for MAH to report serious and non-serious reactions to EMA (29). This was an updated to the previous legislation that only stated as mandatory to report serious ADR.

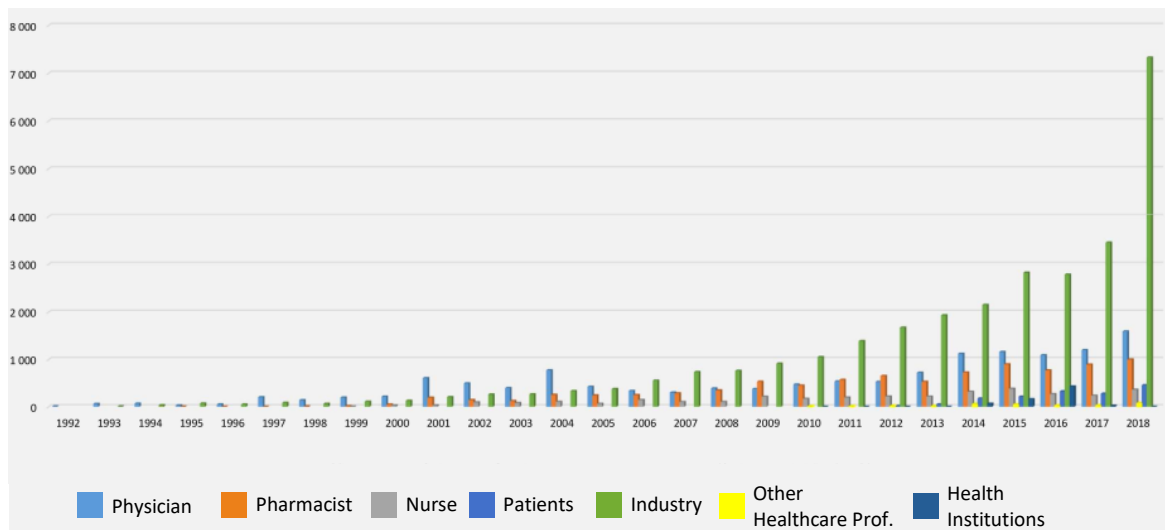


Figure 2 Quantity of spontaneous reports in Portugal per origin, adapted from Infarmed I.P. (30)

The picture also depicts a scenery where the patients are one of the lowest. The values for the US are somewhat different, being the pharmacist the first, followed by physician with industry unable to report (30). According to some numbers from eudravigilance,, Portugal has the lowest rates in Europe for patient reporting, having only 51 reports per million of population. Netherlands in the first place has 706 (31). Overall, Portugal in 2016 had a total of 5690 notifications (517 per million habitants) and the Netherlands had a total number of notifications around the 24000 (1441 per million habitants), excluding the pharmaceutical industry (32,33). These graphs and numbers

only prove that is still a lot to be done to promote ADR notification in patients and healthcare professionals alike.

### 1.3. The causality assessment support tool

As the INFARMED, I.P. as a whole, the Porto Pharmacovigilance u is also committed to promote and maximize the number of suspected ADR reports as much as possible. In order to try to ultimately fulfil this issue, a system was conceived inside our research group. The tool's purpose is to deliver a mechanism of evaluating reports more efficiently, providing a faster and more comprehensive response to the reporter and reinforcing that the time invested into notifying was well spent and the report is being looked into. This is especially pertinent, since fulfilling all the paperwork and the requirements of looking into every website seeking useful information every time a report is received is very time consuming. This is especially alarming, if we take into account that:

- a) the regional offices receive several ADR reports per day;
- b) they have an objective of improving these numbers significantly;

In sum, the current method of evaluating reports is not a scalable procedure and will grow as a problem as the number of reports increase.

All these issues were planned to be mitigated by this system that retrieves the information from the mentioned sites and returns the information in a structured format. A system of this nature may have another benefit, since it was designed to be able to collect and aggregate information from very different providers, it is possible to include documents of very different structures, sources and information in the same database with harmony. Therefore, this system could be used not only for pharmacovigilance centres, but also for several other groups, such as healthcare professionals, pharmaceutical industry and researchers.

Currently, the system only functional feature is to collect information from all the mentioned sources and showing it to the user through a web user-interface. The process starts with the system prompting the user for a brand name or active ingredient of the drug to search and the adverse reaction intended. After receiving the user input, the system searches the abovementioned websites and returns the results. To make this possible, the system is already able to match brand names into generic products and the other way around. Moreover, the translation of the Portuguese terms into English/international terms is currently in progress. This feature, is especially important since it provides a common ground for grouping several information sources and establishes a method for harmonization of information, which is specially impactful since drug and pharmacovigilance information sources vary deeply of the format of information and data available (34).

Information gathering is made through *webscraping*. This option was supported since the websites currently do not have Application Programming Interface (API) and this option was the one who seemed more efficient at the time. It provides a very customizable approach, but with more development intensive requirements. The tool is based on webdriver which mimics a browser instance and repeats the commands stated in the code. This is particularly valuable when the websites have very different methods of showing information and there is a need to reduce the vulnerability of the built code to the target Hyper Text Markdown Language (HTML) structure changes. Regarding the structure of the information provided, the selected sites are very distinct:

- **VigiAccess** includes all the reported suspected ADR categorized by: reaction type (coded with MedDRA), geographic distribution, age group distribution (distributed by age group with focus on younger years and grouping together greater numbers of people as age goes up), patient gender distribution and number of reported suspected ADR per year (from 1969 to present). All of these categories cannot be interlinked and filtered between them. Moreover, the search is made by brand name or by active ingredient and results are given by active ingredient.
- **MHRA** returns information regarding the gender of the patient, age group in groups of 10s, year that a certain ADR report was received back to 1993, type of reporter, entity that received the report, route of administration of the medicine, seriousness and system organ class affected (structured by MedDRA). All the mentioned topics can be filtered. The search is made through active ingredient.
- **Medscape** does not offer information of suspected ADR, but its information of adverse drug effects can have a positive influence on the quality of the information provided. It contains a description of known side effects, and its frequency, without any more information related to it. These are not coded with MedDRA and search can be made with active ingredient or brand names.
- **EMA's** ADR report can be searched with brand or active ingredient, having different tabs for each. The information provided is filtered by age groups (special focus on the 0-18 years and not much distinction afterwards), patient gender, and geographic region from where the notification came (EEA or non-EEA). As opposed to the previous sites, and even though it is not possible for the user to filter most information, several tabs already have that information pre-filtered, such as reaction groups (coded with MedDRA) by patient age group, gender, reporter type and geographic origin. In addition, it contains the opposite as well. Patient age group, gender, reporter type and geographic origin by selected reaction group. Finally, it contains the patient age group, gender, reporter type and outcome per selected reaction (single instance of a reaction group). It uses a system based on Oracle® Business intelligence tool and flash to show the information.

The drugs database was built according the INFARMED, I.P. database and the synonyms were linked by the drug databank and Wikipedia. This is a semi-automatic method and due to that, still error prone. As for additional features, the system can handle user registration for access control. It is possible for a registered user to manipulate the translation of the system and create a dump of information update. Email is also operational in order to receive the documentation in the desired webmail address. Images of the system can be seen in the Annex A.

This will be the system evaluated during this thesis and for which the results are aimed to, not only in a matter of reporting its current state, but also documenting features that can be used in real-world and performance metrics. Another focus will be

the documentation and examination of propositions for additions, changes and improvements of the system itself.

#### 1.4. Pharmacovigilance Informatics: state of the art

This subchapter will act as a primer of the state of the art of informatics applied to pharmacovigilance. What is being done throughout the world according to the existent bibliography regarding data and information system for usage in the pharmacovigilance workflow.

Since the 60s, the requirements, tools and methods used in pharmacovigilance have matured into a well-developed skill set. Nevertheless, and despite these evolutions the healthcare setting has been evolving as well, and pharmacovigilance sometimes shows some difficulty to keep up with the dramatic changes in the healthcare environments of the last 20 years. Whether by the rampantly increase of technology surrounding healthcare professionals, or the new discoveries science has provided in the meanwhile, pharmacovigilance is struggling to keep up and increase its performance accordingly (35). A perfect example of this issue are the websites seen above, where some of them use Flash which is a technology deprecated a long time ago, being that even the browser Google Chrome – which almost 67% usage worldwide (36), will not provide flash support starting 2020. A true interconnection between informatics and pharmacovigilance is shockingly in due. That said, it is important to reflect on how pharmacovigilance as a discipline, and the regional and central offices of pharmacovigilance must change to adapt to this brave new world.

There are some influences that impact developments within the field of pharmacovigilance that help to characterize the current state of this science and the environment in which it is being practiced. One is the increasing awareness, interest and scholarship, since this has been a growing subject in the last 20 years, according to Paul Beninger and Michael A. Ibara, publications related to pharmacovigilance are growing logarithmically (35). Also regulatory developments have been fundamental to the current state, since health and regulatory agencies have pushed for a more regulated format of electronic submission of regulatory documents, including pharmacovigilance reporting for medical products in development and following approval for marketing (37,38). However, Health Authorities are slow to develop regulatory structures and have developed their own separate approaches, with no apparent evidence of harmonization or even the development of collaborative standards across jurisdictions.

Globalization and free market have also impacted this sector, imposing new rules for tenders which can be easily accessed by everyone in the world, providing a framework for outsourcing pharmacovigilance activities (35). Finally trained talent has been an issue over the last few years, since the pharmaceutical industry has recognized a shortage of specialized personnel, creating the innovative medicine initiative (39), which is a public–private partnership that provides academic infrastructure for the education and training of highly skilled researchers, including pharmacovigilance as well, in order to provide generally recognized certifications for pharmacovigilance employment, like the one in the USA.

These are the principal issues affecting the pharmacovigilance industry and the ones expected to impact the most the new technological innovations brought into scene. The words big data, analytics and Artificial Intelligence (AI) are nowadays pretty much worn out buzzwords in the general technological fields and have set the stage for entirely new

operating models that have been emerging in healthcare. These new ways of thinking and using data are only now getting a little traction on the pharmacovigilance scene, so it is expected in the short term to see blockchain, AI, database harmonization and wearables get into play as can be seen below.

**Massively Large Linked Data** could lead to promising results. Teams are employing efforts on gathering information from the most diverse sources, being the EHRs the firsts to be taken into account due to the theoretical massive information these would contain (40). Although yielding good results, these are still very shy attempts and very localized. There is still much to do. More recently, datasets that focus on the combination of electronic health records and social media have been mined for adverse events.

There is some evidence that linking large amounts of data together, far larger than has been done previously, can lead to better adverse-events detection (41). Whereas in the past, only a few types of data have been combined, we can now expect to see many more varied types of data and vastly larger amounts of data being linked together like the ones on figure below. The increasingly better tools of data visualization and intelligence data analysis will be able to mine these more varied and thorough information pools in order to create better knowledge and evidence. This will be crucial for pharmacovigilance practitioners to understand.

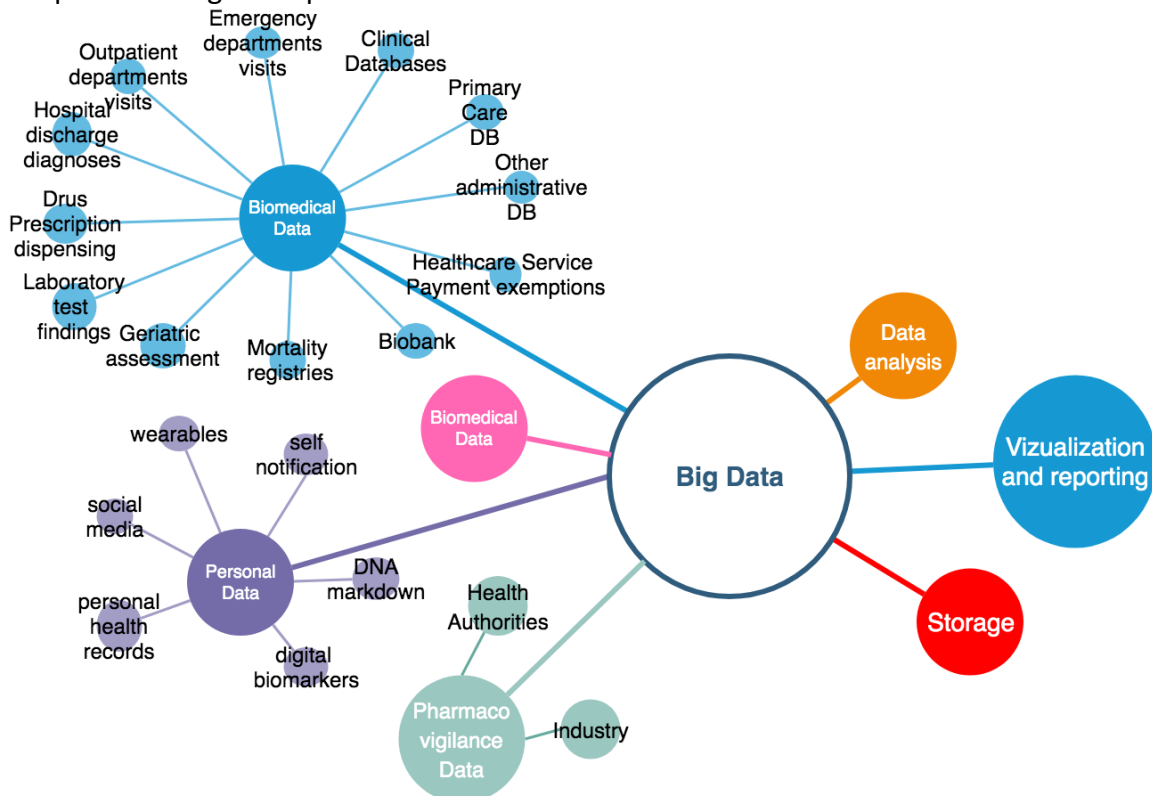


Figure 3 Potential of different data to be linked together

**Task automation** has been an area that gained focus alongside the growth of Artificial Intelligence (AI). AI has supported the appearance of hundreds and hundreds of new applications for the most diverse areas and sectors. So, it was only a matter of time before it reached pharmacovigilance. According to an article of Juergen Schmider et al (42), there are a lot of steps that can be automated or partially automated, which can be seen in the figure below, taken from Schmider et al (42).

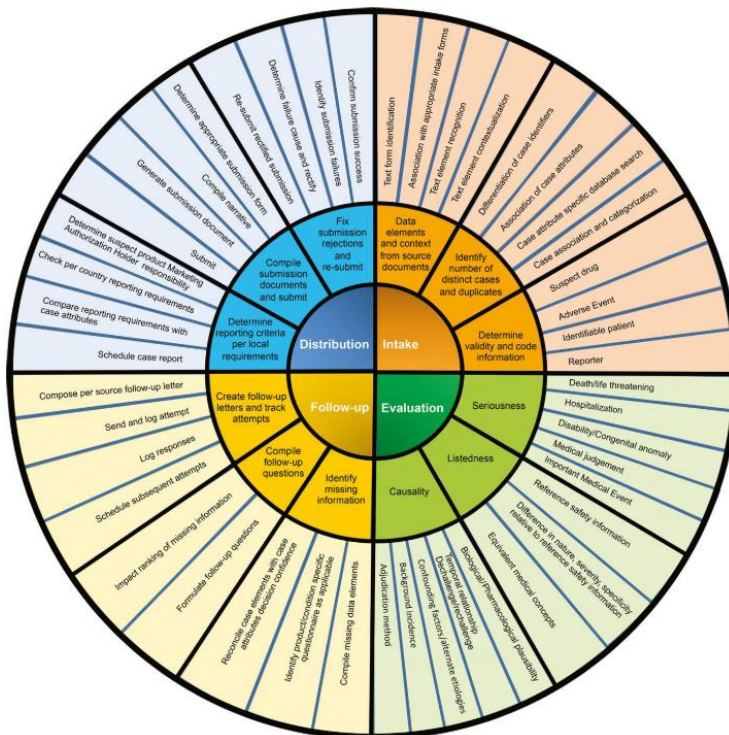


Figure 4 Possibilities of task automation in pharmacovigilance from Friedman et al. (41)

The most appliances seen right now focus of evaluation and intake. There are already a few studies reporting usage of AI on automatic detection of signals on social media (43,44) and EHR (45). Furthermore, there are a few papers stating the capability of employing natural language processing on biomedical literature in order to create a larger database of information (46–48). These can be distinguished into two categories, processing papers in order to find new evidence of ADR reports in clinical trials publications, or finding new drug-drug interactions for further evidence compilation. Although these are only prototypes and proofs of concept, the capabilities of automating time consuming and resource demanding task in the ADR report processing.

**Blockchain** was recently a hot topic due to the price cap of Bitcoin. However, whereas Bitcoin is a financial asset, the technology behind it, the blockchain, can be used far beyond the FinTech world. Blockchain is a type of broadly distributed database that stores a permanent and tamper-proof ledger of data (49). As of 2018, there were found 47 papers stating the usage of blockchain in healthcare. Since this is a frankly new hot topic and technologically demanding, it is a high value (49). The impact is diverse and potentially huge. In a time where counter faction of drugs and data privacy concerns are the highest ever, these could be an interesting take and pharmacovigilance industry must prepare itself. Current applications and proposal affect the areas described in the image below.

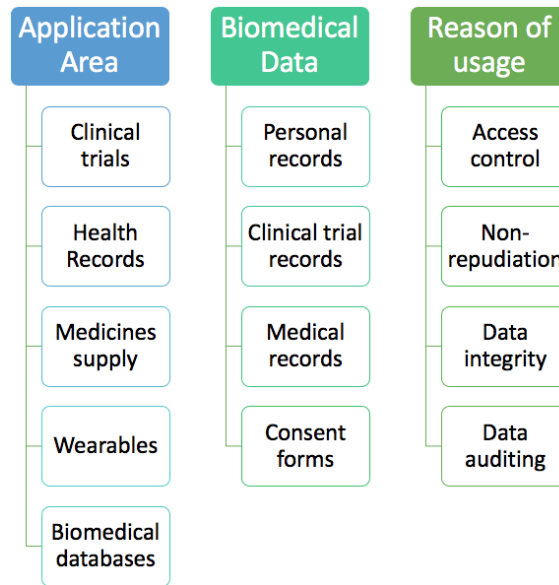


Figure 5 Potential Uses of blockchain in healthcare

In conclusion, there is still a lot of work for providing the necessary tools of effective pharmacovigilance screening of the data currently available. There is a need for more quantity and more interoperability. A primer should be making the data findable, accessible, interoperable, and reusable (FAIR) as well as adhering to international guidelines on transparency and reproducibility (50). As technology evolves and data sources, quality and quantity of information increase, an effort should be made the pharmacovigilance industry to keep up with it, fully embrace technology as a tool and prepare the human resources in the field to be prepared for it.

### 1.5. Objectives

For this thesis, I aim to carry on the developments of our research group of a causality assessment support tool. The main objective is to assess the current state of the tool, interviewing experts and users for understanding the most useful performance indicators and assess those indicators.

Furthermore, it is intended to collect the experts' insights and existent bibliography to create a roadmap for the further developments of the tool, with a focus of facilitating the causality assessment process and in last instance, try to enhance notification numbers.

## 2. Assessment of biomedical systems

The American Medical Informatics Association defines biomedical informatics “as the interdisciplinary field that studies and pursues the effective uses of biomedical data, information, and knowledge for scientific inquiry, problem solving, and decision making, motivated by efforts to improve human health”(35). Over the past several years, the field of biomedical informatics has been successful in establishing definitions and principles that have coalesced into a coherent approach to healthcare practice and data (51).

Nevertheless, even nowadays where healthcare related technology is sprouting like never before, healthcare technology and biomedical systems seem to be the “poor parent” of evaluation. Electronic health systems are a must-have in every hospital, ranging from monolithic systems that cover all activity to highly personalized and specific systems in every clinical service and subservice. Medical devices , which are products, services or solutions that prevent, diagnose, monitor, treat and care for human beings by physical means (52) are booming. Added to this, MedTech Europe, the European trade association representing the medical technology industries, reports that there are currently more than 500 thousands medical technologies available in hospitals, community-care settings and at home (52), being one of the areas that fill more new patents per year, being filled more than 12 thousand requirements in 2017, more than double the ones fulfilled by the entire pharmaceutical industry.

But the biggest boom seen across health-related tech was in the wearable and mobile application’s department. Consumer-grade wearables (which are not medical devices) are currently at 33% adoption (53) (from 9% in 2014). Health related mobile application or “apps” claiming to provide some sort of help, support or benefit to the health of the user, according with data from April 2019 indicates that global medical apps reached a total value of 400 million download in 2018 only – a 15% increase facing 2017. Furthermore, consumer spending in health and fitness apps also tripled during this time (54). Accenture (53) and IQVIA (55) report that 48% of surveyed had at least one health app on their phones or tablets. This value is complementary to the fact that are currently more than 400 thousand “apps” on the different “app’s” marketplaces.

These are signs that the population are embracing healthcare apps. But are these systems able to provide the support they announce? This kind of software are not regulated like medicine and medical devices and have a much more liberal connection to the market and consumer. According to the previous studies, from 2007 to 2017 there were only 570 studies done regarding efficacy of the health and wellbeing and almost half of these were based only on observational studies (55). These means there is almost no evidence the apps are due to provide the benefits advertised and the available evidence is not very well sustained. But there are other issues affecting the evidence of the quality of apps available. There are no regulatory obligations to insert these kinds of systems into the market. All of the health market is very strict and regulated, from drug market approval, medical procedures and/or medical devices, but astoundingly, there is almost none regarding medical apps. This lack of regulation and information to the general public, according to some experts, is possibly harming the adoption of proper

medical apps, instead of promoting it, since a clear and proper regulation can provide the public with a sense of security and confidence (56).

Food and Drug Administration (FDA) has released some guidelines regarding this, but only covers what are classified as moderate-risk (Class II) and high-risk (Class III) mobile medical apps and medical devices, not mentioning the remaining (57). In 2018, the United Kingdom has also released some guidelines regarding this matter, approaching topics of how to classify apps, how to report malfunctions/error prone systems on the yellow card framework, implementing the beginnings of a post-market evaluation for medical apps (58).

All things considered, evaluation of medical systems already has the foundations for health technology established, but has still a long way to go regarding minor applications. Health technology assessment is currently in demand in order to fully implement a need and general assumption to look into health technology and medical devices and gather evidence to support their actions.

### 2.1. What is evaluation?

Evaluation is a term that suffered some changes through the years and is systematically associated with badly noted concepts. But the truth is we evaluate things, situations and thoughts every moment of our lives. As time went by, there's been some tentative definitions of evaluation like the broad definition of "attributing value to an intervention by gathering reliable and valid information about it in a systematic way and by making comparisons, for the purposes of making more informed decisions or understanding causal mechanisms or general principles" (59). More precise concepts for medical informatics appeared in the meantime, like the one of AmmenWerth et al, according to whom, we can apprehend evaluation in medical informatics as "the act of measuring or exploring properties of a health information system (in planning, development, implementation, or operation), the result of which informs a decision to be made concerning that system in a specific context" (60).

The need for effective evaluation of biomedical systems is actually increasingly important in medical informatics since nowadays there is an immense range of different biomedical systems already in clinical practice or aiming to be in the short time. But just like a drug, a system does not automatically correspond to what is intended to do. Furthermore, a system may go through several and strong tests throughout the development phases, but these do not assure that the system will perform as planned in real-life cases with real-life users, as has been documented for numerous cases (61). This has implications on the purpose and efficacy of a system and most of all, carry potential hazards for patient safety (62,63). But these are not the only reasons to justify a thorough evaluation of a system. According to Jeremy C. Wyatt (64), there are several other reasons to perform an evaluation, being:

**Promotional**, in order to encourage the use of the systems (We must be able to reassure physicians that the systems are safe and benefit both patients and institutions)

**Scholarly**, to drive the scientific field of biomedical information to new heights

**Pragmatic**, to help developers learn what the systems is capable of and learn of their features and misuse.

**Ethical**, so the board of a healthcare institution can safely assess which systems are more in tune with their needs, according to

**Medicolegal**, towards reducing liability for developers and clinicians, enabling a safe association of the system and its use.

So, this kind of practices and experiments, not only help provide valuable feedback on the current capabilities of the system at hand, but also help to offer insight about the course the development, biomedical systems should take, driving innovation.

The great range of evaluation purposes and applicability is almost axiomatic. And it becomes evident as well that a proper evaluation is a hard task on itself, since the correct method must be implemented in order to achieve maximum usefulness and results. The range of what can be evaluated is very wide, but there are five major things that can be assessed in an information resource, such as (64):

1. The need for the resource: Evaluators aim to study the clinical *status quo* absent the resource. They determine the nature of the problems that the resource is intended to address and the frequency with which these problems arise.
2. The development process: Evaluators study the skills of the development team and the methodologies employed to understand whether the design is likely to be sound.
3. The resource's intrinsic structure: Evaluators study specifications, flowcharts, program codes, and other representations of the resource that they can inspect without running the program.
4. The resource's functions: Evaluators study how the resource performs when it is used.
5. The resource's effects: Evaluators study not the resource itself but rather its influence on users, patients, and health care organizations.

In order to achieve these assessment results an evaluation process must start by defining the stage of the system for selecting proper evaluation. There are 4 main stages:

- Early stage and problem definition, deciding whether to or/and how to develop a system;
- Pilot testing, in order to support system development;
- Summative test, to check whether the system developed showed benefits;
- Post-marketing, to describe problems and unexpected events occurred during real-world usage.

Consequently, the type of study to apply should be selected, since there are several types, each one intended for different purposes, having different focus and objectives. They are listed in the table below (65), indicating what should be studied and example question that apply to them.

*Table 1 Different Study Types*

Study	Objective	Example question
<b>Needs assessment</b>	Clarify the problem	What is the problem?
<b>Design validation</b>	Evaluate development processed and design	Are the development methods compliant with the best practices?
<b>Structure validation</b>	Expert opinion on the initial form of the resource	Is the resource correctly designed?

<b>Usability test</b>	Appropriateness of the resource to the capabilities of the end-users	Can the intended users navigate appropriately the resource?
<b>Laboratory function study</b>	Performance evaluation	Does the resource have potential to be beneficial?
<b>Field function study</b>	Pseudo-deployment	Does the resource have potential to be beneficial in real world?
<b>Laboratory user effect study</b>	Evaluation of the capability of the end-users to solve problem with the resource	Is the resource likely to change the user's behaviour?
<b>Field user effect study</b>	Behavioural evaluation	Does the resource change users' behaviours in a positive way?
<b>Problem impact study</b>	Evaluation of the mitigation of the original problem	Does the resource have a positive impact in the original problem?

Needs assessment relates to a study whose purpose is to enlighten the problem intended to solve. These studies usually take place before the product/resource is made and takes place in the setting the problem exists or the resource should be employed. The potential users should be studied while they work in real-life problems and tasks in order to understand inadequate information flows or issues during these processes.

The Design validation still relates to a hypothetical resource and/or procedure. In this case, the work focused on reviewing the quality of the processes of information resource design and development. This includes a scrutiny on the theoretical processes and workflows that should be implemented in the resource to be.

Structure validation study steps up and already takes into account a viable prototype in the like of a "minimum viable product" (MVP). The study should be done by an expert with access to the full documentation of the resource and its architecture. The evaluation can range from algorithm implementation to knowledge bases.

Usability testing is intended to evaluate if the users of the system can properly operate the resource. This is applied to understand if an evidence-based and scientifically developed system can in fact help the users to solve the problem initially set to solve. Through the deployment in a controlled environment, it is possible to assess possible improvements to the usability. This is done by presenting the resource to the user and let them navigate while commenting on it or providing a small set of tasks to be completed.

Laboratory function studies takes usability a step further focusing on what the resource can do regarding its purpose. It should tackle issues like the level of performance in the context of it should achieve for the user. Since the input data is key in the study, it should be implemented in the most real-world setting possible.

Field Function studies are very similar to the laboratory mentioned above but is implemented in real world, with real users and real tasks, only missing the access of the final output of the resource to the end-user in order to prevent its actions being affected

by the system. This differentiation is subtle but important since this reflects the maturation state of the resource and a shift of the perspective of evaluating specific task to get the big picture of the possible impact of the resource.

Laboratory-User Effect Studies is a type of study where the actions of a representative subject of the end-user is asked how he would use the system to get to a certain response or output to solve a mock case. This is somewhat similar to usability test but on a grander scale, since the usability focus on if the subject can use the system (i.e., which button to press to get somewhere), but laboratory-user effect focus on if the subject can properly manipulate the resource for getting a response that is meaningful in a certain context.

Field-User Effect Studies is a study that should be implemented in a live scenario, so the resource should be able to perform accordingly. The purpose of this test is to evaluate if the resource is being used by the target population, if the results are in the range of the expected or how the resource is shaping the actions of the users. This is particularly important since there are no guarantees that a potential resource is actually being helpful, or even worse, if its output is not being ill-used and actually decreasing the quality of healthcare.

Finally, there is the Problem Impact Studies, which is the ultimate study since it will focus on what initiated the development of the resource – a problem needed to be solved. Regarding field user effect study and impact study, the difference is similar to the one between efficiency and efficacy. Even if the resource is able to solve the problem, there is no guarantee that the problem will be solved by it, since there is a great amount of environment variable to be taken into account and very difficult to simulate. So, this study type aims to measure the impact (positive or negative) on the initial problem and seeks to go further than the others and get an overall picture of the deployment of the resource such as care processes, cost, team functions and cost effectiveness.

After selecting the study type and its generic purpose, one should decide which kind of evaluation methods should be employed along the study.

## 2.2. Evaluation Methods

Buckley and Chiang define research methodology as “a strategy or architectural design by which the researcher maps out an approach to problem-finding or problem-solving”(66). According to Crotty, research methodology is a comprehensive strategy ‘that silhouettes our choice and use of specific methods relating them to the anticipated outcomes,(67) but the choice of research methodology is based upon the type and features of the research problem (68).

As was stated above, there are a lot of options regarding what should be studied and simple templates regarding what kind of study should be conducted, taken into account what is intended to be studied. But the range of choices regarding studies do not end here. When what will be studied is defined, it is important to deliberate how it should be studied and this is intrinsically connected with the evaluation methods used.

There are several types of classification and categorization of evaluation methods and studies, but one of the most used and accepted is the one developed in 1980 by Ernest House (69) which is supported by philosophical models. These are mainly categorized into two main approaches, objectivist and subjectivist. Each one has 4 subcategories that will be explained in detail in the next chapters (65). For now, the important

distinction is between these two macro classifications. Do note that these study methods are a different type of classification of the study type seen above. Moreover, the choice of study type typically does not constrain the evaluation methods that can be used to collect and analyse data (65).

The major differences are explained by the philosophies behind these two types of studies. While objectivist studies derive from logical-positivist orientation, which preaches that only statements verifiable through empirical observation are meaningful. On the other hand, subjectivist studies follow an intuitionist-pluralist line of thought. This position states that an observation depends largely on the observer, being that two different evaluators can potentially reach two correct yet different conclusions. These two drastic views on evaluation are the corner stones of the evaluation workflow as explained below.

### 2.2.1. Quantitative studies

Quantitative studies or objectivist studies are the most known format of study, that relies only on empirical data. The process is somewhat familiar to a “standard” study and is as illustrated on the figure below.

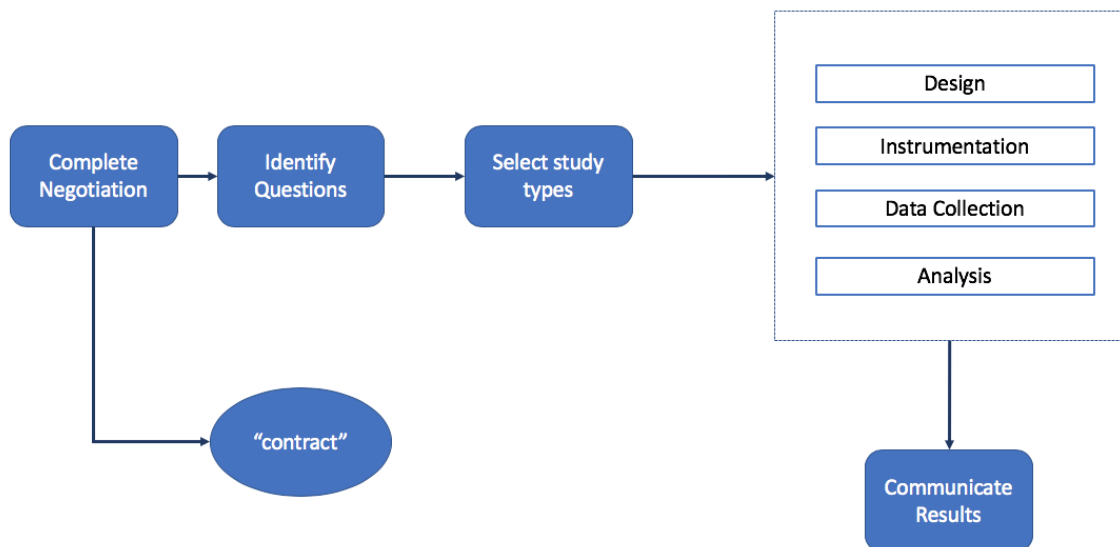


Figure 6 Objectivist study model

In these studies, the workflow is one-way only, being a sequential model of steps. The steps consist on constructing a detailed design of the study, development of instruments for collection of quantitative data, collection and subsequent analysis of these data. In principle, at least, the investigators go through this investigative sequence once and proceed from there to report their findings. The preliminary steps and negotiating a goal for the study, identifying clear and precise question that are intended to be answered and selection of study type for answering these questions.

Since data and objective quantifications take a central role on this kind of studies, the process of measuring data has several key points associated with it. Measurement is assumed as the process of assigning a value corresponding to the presence, absence, or degree of a specific attribute in a specific object (64). Measurement is core since almost all objectivist studies are mainly problems of measurement. This can result in:

- Assignment of a numerical score representing the extent to which the attribute of interest is present in the object;

- Assignment of an object to a specific category.

In these scenarios, object is the entity for which the measurement is done or whose characteristics are being described. As for attribute, we can understand it as a certain characteristic of the object, what is being measured. The attribute can be defined as a very specific concept, such as blood pressure, weight, number of patients seen per day, number of returns the Emergency Room within a certain period. On the other hand, some attributes are more ethereal abstract concept, usually designed by the research team in the specific context of the investigation, such as computer literacy. For this reason, they are recurrently called constructs. Instrument is understood as the technology used to measure the attribute. A measurement is executed through the collections of several observations in order to estimate the true value of the attribute. So, an observation is an act that results in the estimation of the value of the attribute, with the application of an instrument. The employment of multiple observations is done, because it allows for the determination of how much variability exists across observations, which is necessary to estimate the error inherent in the measurement. The measurement tools and objects must be well defined before starting the study, especially in the field of medical informatics, as there are a lot of grey areas regarding proven instruments of measure or well accepted variables for measurement.

There are several types of measurement techniques and formulas, since attributes are very distinct in the way they are represented and expressed, such as discrete or continuous values and so on. Measurement of a nominal attribute can be explained as the adjudication of an object to a certain category. These categories do not have strict connections and order. As for ordinal level of measurements, again focus of allocating an object to a category, but this time categories have order or ranking. These two types of measurements often fall under the nomenclature of categorical variables. Interval level measurements assign the attribute to a continuous numerical value with an arbitrary chosen zero point. Typical examples are Fahrenheit and Celsius scales, since the interval between degrees of temperature have the same significance along the scale. Ratio measurements are similar to the interval measurements, with the added feature of the zero being a true point, enabling the possibility of assigning meaning to the ratio of two measurement results in addition to the difference between them, because the divisions between the points on the scale have an equivalent distance between them, as well.

Having established the foundations of concepts for objectivist studies, it is time to dwell on the type of categories of evaluation method existent. Since the measurements are key in these studies, there is a clear distinction of two main studies objectives. The most common are demonstration studies, which address questions of substantive and practical concern to the stakeholders for an evaluation study. But they would be less than useful without measurements studies that seek to determine how accurately and precisely an attribute can be measured in a population of objects. Measurements studies are particularly important since they develop and refine methods for making measurements that will be used in demonstration studies, establishing a link between them like the image below.

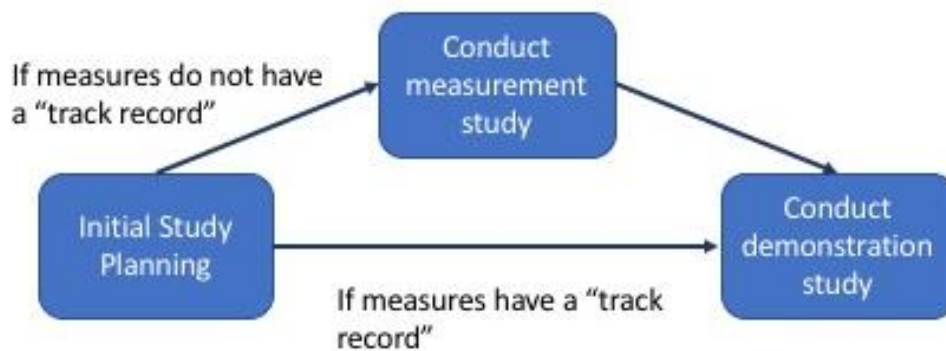


Figure 7 Relation between measurement and demonstration studies (64)

Demonstration studies themselves and the more associated with quantitative studies, and are divided into three main terms that define different intents. There are descriptive, comparative and correlational studies.

**Descriptive Studies** aim to estimate the value of a dependent variable or set of dependent variables and do not have independent variable. Descriptive research is used to describe characteristics of a population or phenomenon being studied. It does not answer questions about how/when/why the characteristics occurred. Rather it addresses the "what" question (70). Even though this type of studies can sometimes look like an simple exercise, they can be difficult to execute properly and be quite informative (64).

**Comparative Studies** are based on contrasting conditions. The work system of these kind of studies relies on assigning subject into one of the conditions or classifying them into one, based on their characteristics. Then a variable is assessed for the subjects and the results are compared across conditions, since usually the independent variable of interest for the study and the dependent variable was the measured one. By comparison, it is often possible to speculate about cause and effect relationship between dependent and independent variables.

Inside these study types, there are a lot of variants, like the historically controlled experiment, which is used as a synonym of a before-after study. In this study, the investigator measures a baseline for a certain variable and then makes the same measurement afterwards. For example, a researcher makes a measurement of medication duplication before a pharmacy information system is installed, and then make the very same measurement afterwards. This study has a few downfalls, since the alteration in the measurements can be result of a lot of difficult to assess variables.

For this reason, controls were added to studies, in order to overcomes some of the before-after flaws. For this, simultaneous controls are employed, where a control group is used for simultaneous evaluation as the intervention group. In the above example, it would be added a measurement for pharmacist not using the newly implemented Pharmacy Information System, resulting into stronger evidence. There are main two types as well – randomized and nonrandomized.

A non-randomized design is used when it is difficult to realize a randomization of participants into groups for comparison or it is an ethical challenge to do so. The design can involve the use of prospective (after) and/or retrospective (before) data from the same or different participants as the control group, being that measurements before and after help to further eliminate other change taking place into the research environment (71,72).

In randomized controls, the participants are randomly assigned to two or more groups. The design is prospective in nature since the groups are assigned at the time of the study, after which the intervention is applied then measured and compared (73).

**Correlational studies** are meant to test a hypothesis of relationship among a set of variables. The independent variables are the hypothesized predictors of an outcome of interest, which is the dependent variable. They aim to evaluate how related two or more variables are (74). They are mainly used to evaluate event exposure, prevalence of diseases or risk factor into a target population.

Commonly known correlational studies are Cohort studies, where a group of subjects is observed over a period of time. Those exposed and not exposed to the are compared for differences in one or more predefined outcomes. Cohort studies may be prospective where subjects are followed for a time period into the future or retrospective for a period of time in the past. Comparisons are usually evaluated at the beginning of the study as baseline, then repeated over time at preestablished intervals for differences and trends.

There are also cross-sectional studies, which are considered a type of cohort study where only one comparison is made between exposed and unexposed subject. They enable researcher to take a snapshot of the outcome and the associated characteristics of the cohort at a specific point in time.

Finally, there are case-control studies, where subjects in a group that was exposed to the event are matched with those not exposed. The groups are then compared for differences in some pre-defined outcomes. Case-control studies are retrospective since subjects that were already exposed to the event are selected then matched with unexposed subjects, using historical cases to ensure they have similar characteristics (72).

In all of these studies, there are a few things that are common, like participant selection and therefore, sampling. This is crucial since the scientific reasoning for having a selection of the population target in order to study the overall effect on the population as a whole.

### 2.2.2. Qualitative studies

Qualitative studies or subjectivist studies rely on a completely different set of assumptions than those that provide sustenance for quantitative studies. Besides, it is understood that the knowledge that can be acquired by this kind of studies is different and complementary to the one obtained through objectivist studies. Subjectivist studies are somewhat unfamiliar to anyone that guides its research by the core of the traditional scientific method. Nonetheless, they are equally rigorous methods and have been increasing in performance and acceptance through the years (75) as solo demanded research or in combined form with quantitative.

Four different types of approaches are related to subjectivist studies: - quasi-legal, art criticism, professional review, and responsive/illuminative and are as follow:

- **Quasi-legal** is a type of trial or legal adversary proceeding where a mock-up scenario is played out. The idea is to judge the object of evaluation, where Proponents and opponents of the resource offer testimony and may be examined and cross-examined in a manner resembling standard courtroom procedure. Based on this testimony, a jury witness to the proceeding can then make a decision about the merit of the resource. As in a debate, the issue can be

decided by the persuasive power of rhetoric, as well as the persuasive power of that which is portrayed as fact. There are few published examples of this technique formally applied to informatics, but the technique has been applied to facilitate difficult decisions in other biomedical areas.

- **Art-criticism** approach relies on formal methods of criticism and the principle of “connoisseurship”. Under this approach, an experienced and respected critic, who may or may not be trained in the domain of the resource but has a great deal of experience with resources of this generic type, works with the resource over a period. She or he then writes a review highlighting the benefits and shortcomings of the resource. Within informatics, the art criticism approach may be of limited value if the critic is not expert in the subject-matter domain of the biomedical information resource under review. For example, if the resource provides advice to users or automates a task that was heretofore performed manually, a critic without domain knowledge could offer useful insights about the resource’s general functioning and ease of use, but would be unable to judge whether the automated task was carried out properly or whether the advice provided was clinically or scientifically valid. Because society does not routinely expect critics to agree, the potential lack of interobserver agreement does not invalidate this approach. Although they tend to be more informal and tend to reflect less direct experience with the resource than would be the case in a complete “art criticism” study, software reviews that routinely appear in technical journals and magazines are examples of this approach in common practice.
- **Professional review** approach is the well-known “site visit” approach to evaluation. It employs panels of experienced peers who spend several days in the environment where the resource is deployed. Site visits often are directed by a set of guidelines specific to the type of project under study but sufficiently generic to accord the reviewers a great deal of control over the conduct of any particular visit. They are generally free to speak with whomever they wish and to ask of these individuals whatever they consider important to know. They may request documents for review. Over the course of a site visit, unanticipated issues may emerge. Site visit teams frequently have interim meetings to identify these emergent questions and generate ways to explore them. As a field matures, it becomes possible to articulate formal review criteria that could be the focus of site visits, supporting application of the professional review approach. In biomedical informatics, the evolving evaluation criteria for computer-based patient records is one example of such guidelines.
- **Responsive/illuminative** approach seeks to represent the points of view of those who are users of the resource. Users who are very proximate of the resource operation environment can also be represented. The goal is to understand, or “illuminate,” rather than judge the points of view. The methods used derive largely from systematic studies of people and cultures, given that the investigators submerge themselves in the environment where the resource is operating. Unlike other studies, the design of these studies is

not rigidly predetermined, as they develop dynamically according to the investigator experience and history. The study begins with a minimal set of orienting questions, being that deeper questions that receive in-depth analysis will appear and develop over time. As for data collection, researchers often use observation techniques, interviews or review of documents. The reports are frequently presented as narratives. Many examples of studies using this approach can be found in the literature of biomedical informatics.

Evaluation questions for qualitative studies are usually more intricate than those usually associated with objectivist studies and often require multiple reads to grasp the full scope of the question. It is common as well that several smaller explanatory questions derive from the main question and are viable as well. Additionally, these

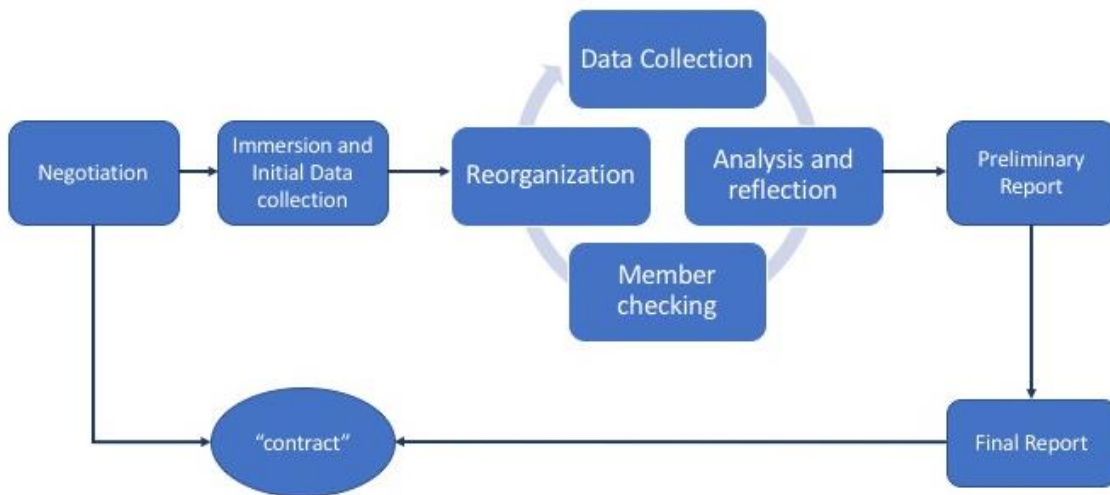


Figure 8 Subjectivist Study model

questions, due to its nature, are often difficult to respond with quantitative studies. This happens because subjectivist studies are aimed to address the points of views of those who are intended to be the users of the system in cause, or are significant player in the environment the resource is expected to act on. It is an important tool to document how the various actors and groups “see” the resource, instead of assuming there is a consensus when there is no reason to believe one exists. The organization of a qualitative studies follows the steps in the image below.

In short, the steps are:

**Negotiation of the “ground rules” of the study**, for arranging an understanding between the investigators and those commissioning the study, if there are any. This is for making sure everyone involved in the early phases of the study are aware of its purpose and aim. It also should be clear for everyone the methods intended to use, sources of information and the general structure of the final report. It works as a contract between all parties.

**Immersion into the environment**, in order to the investigators to get acquainted to the work environment. This can be understood as a kick-off meeting and further initial setups for the investigators. This can include formal and informal introductions and silent presence in distinct meetings and events. It is a way of building trust since openness and trust between scientists and those in the field is essential, since it is

intended that the study disturbs the least possible amount of the environment object of study.

**Initial data collection to focus the questions** is a step difficult to mark since this is a more or less automatic step after the immersion into the environment. Getting to know the people and the object of study will help sharpen the questions and guide the study, but could also shape the investigator's views. Even though this shaping is a natural process, should be taken into account along the study.

**The Iterative loop** is the core of the study, structured on cycles of data collection, analysis, reflection and reorganization. Data is collected on predetermined checkpoints and interpreted in sight of what is known at the time. Reflection proceeds as to contemplate the new findings in each cycle and finally reorganization that appears as a revised plane and agenda supported by the new insights and findings for the next cycle. Even though the image shows a unidirectional arrow in the cycle, counter clockwise movements are also possible and encouraged. This will not mean something wrong happen, but only that clarifications were deemed as needed. An important part is sharing the new findings with the participants, and a major difference from an objectivist study, since this would mean an "unblinding" of the study. Nevertheless, subjectivists studies rely on trust and growingly informed participants and are seen as key.

**Preliminary report**, refers to the first draft version of the final report that itself is used as a tool and investigation instrument. The more a document is shared, the more validity of the findings can be obtained. The overall reaction to a report can be enlightening and helpful for the final study findings and results.

**Final report**, should be distributed as originally negotiated. In subjectivist evaluation studies, distribution of the report is often accompanied by "meet-and-greet" sessions that allow interested persons to explore the study findings interactively and in greater depth.

In sum, qualitative studies are a less known type of study but equally useful and powerful for answering scientific question and evaluate biomedical systems.

# 3.Design & Methods

Regarding the study protocol, it is expected to go along the lines of the following steps:

- 1) Defining question
- 2) Design study
- 3) Instrumentation
- 4) Data collection
- 5) Data analysis
- 6) Report

The defining questions will be:

- **What capabilities should a causality assessment support system have in order to improve report numbers?**
- **Is the current system able to act upon expected?**

For the design study it will be conducted based on a mixture of quantitative and qualitative methods. The qualitative methods will be aimed to gather in-depth viewpoints from Key Opinion Leaders (KOL) in the areas of pharmacovigilance and medical information systems development. Quantitative methods will act as a way of factually describe the system metrics, selected through the expert's opinions and experience.

This coupling of methodologies aim is to bring together the keenness of measuring pervasiveness of "known" phenomena and central patterns of association of the quantitative evaluation and the ability for identification of previously unknown processes, explanations of why and how phenomena occur, and the range of their effects of the qualitative evaluation (76,77). This mixed methods evaluation, then, is more than simply collecting hard factual evidence, and more than just collect opinions from experts. It involves the intentional collection of both quantitative and qualitative data and the combination of the strengths of each to answer the research questions (78). More concretely, data collection should follow by interviewing experts to debate the current problems of causality assessment, ADR reporter engagement and how to improve report ratios. Afterwards, technical and functional analysis will be made to the tool, in order to document its state and features. This will be largely influenced by the interviews, but a pre-set of key performance indicators are already defined. Next will be a though process for delineating and justify the next developments steps of the tool with the purpose of enhancing reporting rates, minimalize automated tasks of the regional pharmacovigilance offices and create more knowledge and value with the data available.

The analysis and report will be conducted along the thesis, based on the evidence gathered and presented in the results section.

## 3.1. What capabilities should a causality assessment support system have in order to improve report numbers.

For this first question, a Needs Assessment will be implemented. This study will be the quarter-stone of the project since it will provide insight into the capabilities the system should have and what are the potential new features that such a system should

have. This study will be conducted through a subjectivist study, collecting the opinion of different players. This type of study is used to seek to clarify the information problem the resource is intended to solve or ameliorate.

That said, one of the main issues to elucidate will be linked to features that a causality assessment support tool can or should have to improve response speed to notifications. Some studies were already made to discuss this issues (26,27) reporting that a faster and more thorough response to notifications can improve the connection with the notifier and act as a positive reinforcement in order for them to notify more and somewhat act as an evangelist of the national ADR notification system. This study will be used to increase the range of possible answer to this question and make the already available ones more robust.

The process for this study will follow largely the steps explained by professors Charles Friedman and Jeremy Wyatt on the book Evaluation methods in Biomedical Informatics (64) and the premises discussed above. The contract is based on this very text, where the general aim of the study is explained, specially the question that I am trying to answer. It contains also the kinds of methods to be used and which data sources should be utilized.

#### 3.1.1. Research methods

The core of the study will be based on interviews to Key Opinion Leaders (KOL). An interview is a type of framework in which the practices and standards should be not only recorded, but also achieved, challenged and as well as reinforced (79). There are two types of interviews: Formal and Informal

Formal interviews are occasions where both the investigator and interviewee are aware that the answers to questions are being recorded for direct contribution to the evaluation study. On the other hand, informal interviews are more like spontaneous discussions between the investigators and members of a team that occur during routine observation and are also part of the data collection process. Informal interviews are invariably considered a source of important data and should not be dismissed easily.

Formal interviews have very distinct degrees of structure. At one end is the unstructured interview, whose structure is non-existent and there are no previously prepared questions. At the other end is the fully structured interview, based on rigid script of questions that are questioned exactly the same way and always by the same order, being very intolerant to deviations from the pre-defined path.

Between these extremes is the semi structured interview, where the researcher indicates prior to the session a couple of themes that are wished to be addressed. Nevertheless, it is flexible to the order in which these themes are attended, and unlike the structured interview, discussion of topics not on the pre-specified list are welcomed. To accomplish this, interviews usually prepare a guide, which is a schematic representation of question and topics. It is similar to the script of the structured interview, but with added topics and themes plausible to be explored (79). This guide helps the interviewer to explore different interviews more systematically and thoroughly, increasing the optimal use of the interview time. The questions can be similar, having for example a core question with an amount of different associated questions around the core concept.

Before starting each interview, a checklist will be crossed in order to ensure the optimal result. Based on the work of Ranney et al., will be as stated below (80):

1. Use at least two recording devices;
2. Check the room ahead of time;
3. Bring facial tissues;
4. Refreshments;
5. Paper and pencils for participants to write down thoughts;
6. List of resources (as applicable, according to the interview topic);
7. Assure privacy of the room; post a sign on the interview door, or consider using a white noise machine to obscure the discussion;

Regarding the session properly said, it will be recorded unless a respondent refuses to do so. This method allows for the interview data to be captured more effectively. Moreover, hand written notes during the interview are relatively unreliable, and the researcher might miss some key points. The recording of the interview makes it easier for the researcher to focus on the interview content and the verbal prompts and thus enables a generation of a “*verbatim transcript*” of the interview. The session should begin with an introduction that reviews the goals of the study, rules of the interview, an overview of the steps and phases of the session and a confidentiality statement if needed. This introduction is often read as is, but can be paraphrased as long as all pertinent points are discussed. Then, the interview properly said starts with an “icebreaker” question. The Icebreaker is a type of question that eases the interviewee into the flow of the session, being an easy and non-controversial question. The icebreaker is usually followed by topic headings, each of which will contain an opening question, main questions, follow-up questions, and probes (80). Opening questions are for introducing the subject and are very straightforward and non-sensitive, just to introduce the topic and get the interviewee talking. Main questions are the major topic of the research. A more generic question and usually more controversial, being that the reason these are not attended right away and are generally open-ended to avoid brief or yes or no responses. The follow-up questions are more in-depth and meticulous around the major topic and are used to get more information and details within pre-planned areas of inquiry. Probes are short questions intended to be tools for retrieving more context or clarifications (e.g. “Who did this? How did you feel?”). Number of probes are completely subjective and may even not be asked none, depending on whether or not a participant spontaneously discusses the information requested by the probes. (80).

Questions are usually divided according to main theme. Each theme (and, sometimes, each question) will be accompanied by an “intent statement,” which is meant to explain the goals of the question, maximize reproducibility of methods, and ensure that all needed information is gathered. A general rule of thumb regarding interview sequence is that more sensitive or difficult questions should appear later in the interview. This will enable a more established connection and comfort between the interviewer and the participant. One alternative is approaching these questions after the participant’s lead. A skilled moderator will almost always make changes to the order and specific wording of questions. It should be able to depart from the guide as needed to follow interesting and relevant ideas that appear during the session.

Interview sessions should close with the interviewer providing a brief summary of the conversation, allowing a clarification or information addition. If the interview is made within a research group, a debriefing could be useful to have after the participants

leave, but before the research staff departs a session. This allows the moderator and note taker to record initial impressions while the verbal and non-verbal data are fresh in memory, clarify any ambiguous data that were collected, and record any pertinent details of the session that may not be captured in the notes or recordings. If the prospect is made alone, a reflection phase is useful as well, comprehending on a period of time for the researcher/interviewer to annotate any piece of information that he thinks could be handy later on and cannot be discovered in the recordings.

The length and exact content of an interview guide will depend on the type of study, the goals of the study, and the amount of time that is available for the interview session. Additionally, most interviews should last no more than 90 minutes, as longer sessions can become extenuating for participants and interviewer alike and affect the data quality and veracity.

As for the data analysis strategy used the study will depend on the research question and qualitative design. Nevertheless, the central idea for data analysis generated from interview sessions should be based on reviewing the data (in the form of transcripts, audio recordings and/or detailed notes), applying descriptive codes to the data and categorizing codes in order to find patterns. These type of patterns can exist within a single interview or across multiple interviews depending on the research question and design (81).

Regarding presentation, it is common to create narrative that mimic or enable the reader to understand at least a major part of the experience the research team had during the process. This is possible through a detailed description of the interviewee's points of views and opinions and since the participants words are the data, it is not unusual to find a great number of quotes for justifying the findings.

In sum, there are a few key characteristics to that lead to a good research using semi structured interviews and they are (82):

- Loose, flexible structure
- Iterative
- Schedule in advance
- Gathers information from key informants who can inform the topic
- Insight into participant perspectives
- Deep exploration of participant thoughts and experiences

The steps for creating the interview are the ones from table below (82):

*Table 2 Steps for creating successful interviews*

Step	Task
1	Determining the purpose and scope of the study
2	Identifying participants
3	Considering ethical issues
4	Planning logistical aspects
5	Developing the interview guide
6	Establishing trust and rapport
7	Conducting the interview
8	Memo and reflection
9	Analysing the data
10	Demonstrating the trustworthiness of the research

### 3.1.2. Design

The group for interview will be selected according to a heterogeneous perspective of all the different intervenient of the spontaneous ADR reporting and general pharmacovigilance workflows. It is planned to include interviewees from different backgrounds to cover expertise and generic assumptions of everyone involved in the pharmacovigilance system and generic drug circuit. Furthermore, it is also intended to include healthcare providers and patients for covering the ADR reporters. For healthcare providers, it is currently planned to include medical doctors, pharmacists and nurses since these three were the most active healthcare professional according with the 2017 report of INFARMED, I.P. (83). Finally, a specialist in medical software development will be addressed as well to participate with a more technical point-of-view and general evaluation of feasibility of the previous answers.

The target will be chosen based on the roles mentioned above and the research contact network, along with some more formal invitations outside the network. These invitations will be mainly done by e-mail, in person or telephone. As for the interview properly said will be made preferably in person or telephone if the need arises. The ideas to be explored in the interviews are depicted below, being that depending on the interviewee profile and answers, different weights could be given to each question:

- Why are notification number in Portugal so low?
  - How can we improve them?
  - What each one of the players could do/should do in order to improve them?
  - Is the response from official entities satisfactory?
- Could an information system help improve these numbers?
  - What capabilities should an information system have in order to improve these numbers?
  - The current system could be helpful for someone else?
- What could be done to improve feedback to reporters?
- Is the current response of the system satisfactory?
- The proposed roadmap for the system is feasible and satisfactory?
- What characteristics are more suitable to make a system like the one depicted be successful?

These questions will be the core of the semi-structured interview. The interviews are estimated for a duration of 20 minutes to a maximum of 1 hour (84). The recording will be made with a mobile telephone, two when possible to add redundancy. If interviews are by phone, record the call and by skype with a screen capture software.

### 3.2. Is the current system able to act upon expected?

Following the needs assessment, the study will follow a Field Function study in order to evaluate if the system has the potential to be beneficial in the real-world. Since this is an already a semi-deployed system, the purpose will be not only to assess capabilities for current features, but also measure existent capabilities regarding future features. To accomplish this, an objectivist study will be employed, in particular a descriptive study.

The parameters will be discussed in the focus group, but some of the core features will be evaluated:

- Speed of returning results;
- Number of times the search returns an error;
- Number of times the product is not found;
- Number of times the result is not correct;
- Reproducibility of results;
- Possibility to add other sources of information collection;

Finally, after the interviews, if any more Key Performance Indicators (KPI) are enlightened and feasible, will be added to this part of the study.

For this evaluation, the system was evaluated in a setting similar to real-usage. It was measured the time with resource to a chronometer and google Chrome Developer Tools and every test was done at least 3 times and the average was used as metric. Unless for tests aimed for it, if an error was returned, the test would be repeated. In occasions where different sites have distinct responses, it will be pointed out. For the correct results, a drug was selected taking into account it would be accessible on all 4 websites. Then the true results would be registered and then a search on the system was conducted. Error was described as the system returning a result with an error message, timeout was defined as no answer for a period over 5 minutes.

## 4. Results

### 4.1. Interviews results

In total, 6 interviews were done, bringing together viewpoints from different areas of the drug circuit and pharmacovigilance. They all belonged to the great Porto Area. The areas of professional activities were physician, hospital pharmacist, community pharmacist, nurse, clinical software developer and pharmacovigilance officer. The interviews had an average duration of 33 minutes, being the longest a total of 41 and the shortest 15. 5 of the 6 interviews were made in person and one by telephone.

### 4.2. Feature evaluation

For the feature evaluation, a few tools were developed to evaluate the desired parameters more easily and start a performance and metrics evaluator. Since data related to the website information provided was in a Not Only Structured Query Language (NoSQL) database – mongoDB, an extraction was made and the resulting JavaScript Object Notation (JSON) can be easily manipulated with Python. Taken this into account, a tool was developed, receiving as input an mongoDB export and resulting into several metrics seen presented in this paper.

Regarding capability of returning a result, the totality of searches in the system were used. This comprised a set of 330 different searches. From these, 116 unique drugs were found. The results for error returned and drugs not found are the ones like the table below.

Table 3 Performance Indicators of the system

WEBSITE	Error	Drug Not Found
<b>Medscape</b>	23 (7.0%)	104 (31.5%)
<b>MHRA</b>	29 (8.8%)	77 (23.3%)
<b>VIGIACCESS</b>	51 (15.5%)	0 (0%)
<b>EMA</b>	182 (55.2%)	1 (0.3%)

As for individual drugs, only 48 drugs were not found and 61 returned some kind of error. The table below shows the relation of drug to an error and absence of results.

Table 4 Results of non-successful searches in the system

Portuguese	English	Not Found	Error
Abacavir Lamivudina	Abacavir Lamivudina Teva	MS MHRA	VIGI EMA
Acemetacina	Acemetacin	MS	
Aciclovir	Aciclovir	MS MHRA	EMA
Ácido acetilsalicílico	Acetylsalicylic acid	MS MHRA	EMA
Ácido fusídico	Fusidic Acid	MS MHRA	
Ácido zoledrónico	Zoledronic acid	MS MHRA	EMA
Agomelatina	Agomelatine	MS	EMA
Amoxicilina	Amoxicillin	MHRA	EMA
Amoxicilina Ácido clavulânico	Clavamox DT	MS MHRA	VIGI EMA
Atorvastatina	Atorvastatin	MS	
Buprenorfina	Buprenorphine		VIGI

Azilsartan medoxomilo Clorotalidona	Edarclor	MS MHRA EMA	
Captopril			VIGI
Carboplatina	Carboplatin		EMA
Cefazolina	Cefazolin	MHRA	
Ceftriaxona	Ceftriaxone		EMA
Ciproterona	Cyproterone acetate	MS MHRA	VIGI EMA
Citrato de sódio Laurilsulfoacetato de sódio	Microlax	MS MHRA	VIGI EMA
Clonixina	Clonixin	MS MHRA	EMA
Cobamamida	Cobamamide	MS MHRA	EMA
Dabigatrano etexilato	Dabigatran etexilate	MS MHRA	VIGI EMA
Dexmedetomidina	Dexmedetomidine		VIGI
Efavirenz			MHRA
Emtricitabina Tenofovir	Tenofovir disoproxil	MS MHRA	VIGI
Entecavir	Entecavir		VIGI EMA
Entacapona	Entacapone		EMA
Escitalopram	Escitalopram		EMA
Etanercept	Etanercept		EMA
Etinilestradiol Gestodeno	Minigeste	MS MHRA	EMA
Etoposido	Etoposide		VIGI
Evolocumab			VIGI
Factor VIII da coagulação humana Factor de Von Willebrand humano	Wilate 1000	MS MHRA	VIGI EMA
Flucloxacilina	Flucloxacillin	MS	EMA
Furosemida	Furosemide	MS	
Fluticasona	Fluticasone propionate		VIGI
Ginkgo Biloba	Ginkgo biloba	MS MHRA	EMA
Golimumab	Golimumab		EMA
Goserrelina	Goserelin	MS	EMA
Ibuprofeno	Ibuprofen		EMA
Hidroxicarbamida	Hydroxyurea		EMA
Indometacina	Indomethacin		EMA
Infliximab	Infliximab		EMA
Iopromida	Iopromide	MS	
Levofloxacina	Levofloxacin		VIGI EMA
Lasofoxifeno	Lasofoxifene		VIGI
Levofolinato disódico	Levofolic	MS MHRA	
Lornoxicam	Lornoxicam	MS MHRA	
Levotiroxina sódica	Levothyroxine		EMA
Macrogol e outras associações	Vitamin C		VIGI
Melperona	Melperone	MS	
Mirtazapina			MHRA
Nifedipina	Nifedipine	MS	EMA
Nitrofurantoína	Nitrofurantoin		EMA
Nimesulida	Nimesulide	MS MHRA	
Osimertinib	Osimertinib		EMA

Paclitaxel	Paclitaxel	MS	EMA
Oxaliplatina	Oxaliplatin		EMA
Paracetamol	Acetaminophen	MS MHRA	EMA
Paracetamol Cafeína	Ben-u-ron Caff	MS MHRA	VIGI
Peginterferão beta-1a	Peginterferon beta-1a	MS MHRA	
Pertuzumab	Pertuzumab		EMA
Petidina	Pethidine	MS	
Piperacilina Tazobactam	Piperacilina Tazobactam	MS MHRA	VIGI
Rifampicina	Rifampicin	MS	
Rilmenidina	Rilmenidine	MS MHRA	
Rituximab	Rituximab		EMA
Rosuvastatina Ezetimiba	Rosuvastatina Ezetimiba	MS	VIGI
Secucinumab	Secukinumab		EMA
Silodosina	Silodosin	MHRA	
Sulfametoxazol Trimetoprim	Sulfametoxazol Trimetoprim	MS MHRA	VIGI EMA
Tansulosina	Tansulosina	MS MHRA	
Telmisartan Hidroclorotiazida	Actelsar HCT	MS MHRA	
Tenofovir	Tenofovir disoproxil	MS MHRA	VIGI EMA
Tramadol	Tramadol		EMA
Vacina contra a difteria e o tétano	D.T. Vax Adulto	MS MHRA	EMA
Vacina contra a difteria, o tétano, a tosse convulsa, a poliomielite e o haemophilus tipo b	Pentavac		VIGI
Vacina contra a gripe	Vaxigrip		VIGI

For a proper match, it is important to check how many database entries have correspondence with the English term. So, the current data is:

**23702 entries with correspondence (72,62%)**

**8938 entries with no correspondence (27,38%)**

From these values, a collection of pharmaceutical forms that had more problems was evaluated. There was a total of 65 pharmaceutical forms that had no drug associated with a translation (annex C). The number of drugs associated with these pharmaceutical forms was mainly 10 or less. The major pharmaceutical form with a great absence of translation is the injectable suspension (250 of 269 with a lack of translation) and perfusion for emulsion (58 of 66 with a lack of translation).

Finally, concerning results returned, the zero results were evaluated. From the database, 228 groups of drug and reaction returned zero values.

Regarding search time, the tests were done on the 19<sup>th</sup> of August and were tried 3 types of searches:

1. One reaction
2. 3 Reactions from the same group
3. 3 Reactions from distinct groups
4. 5 Reactions from the same group
5. 5 Reactions from distinct groups

The results are the ones shown below.

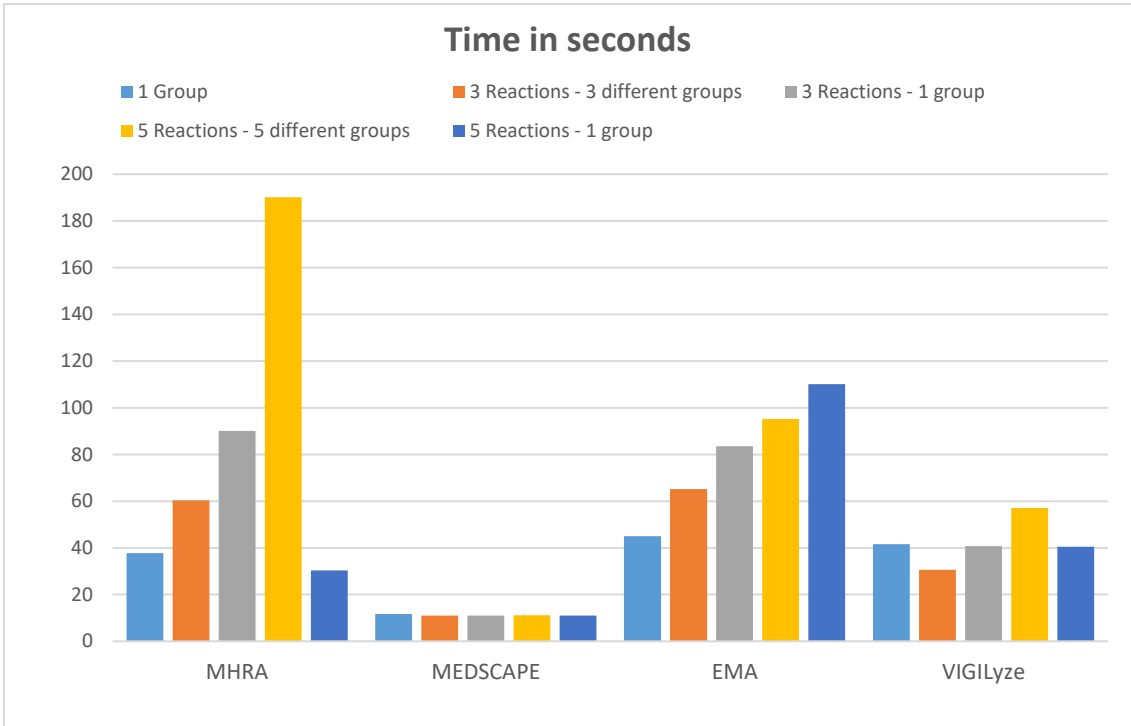


Figure 9 Time of search for different sites and requests

After that, the loading times for the reactions inside each reaction group were also evaluated, resulting into the graph below, that shows the top 10 slower times. The rest of them were all below five seconds.

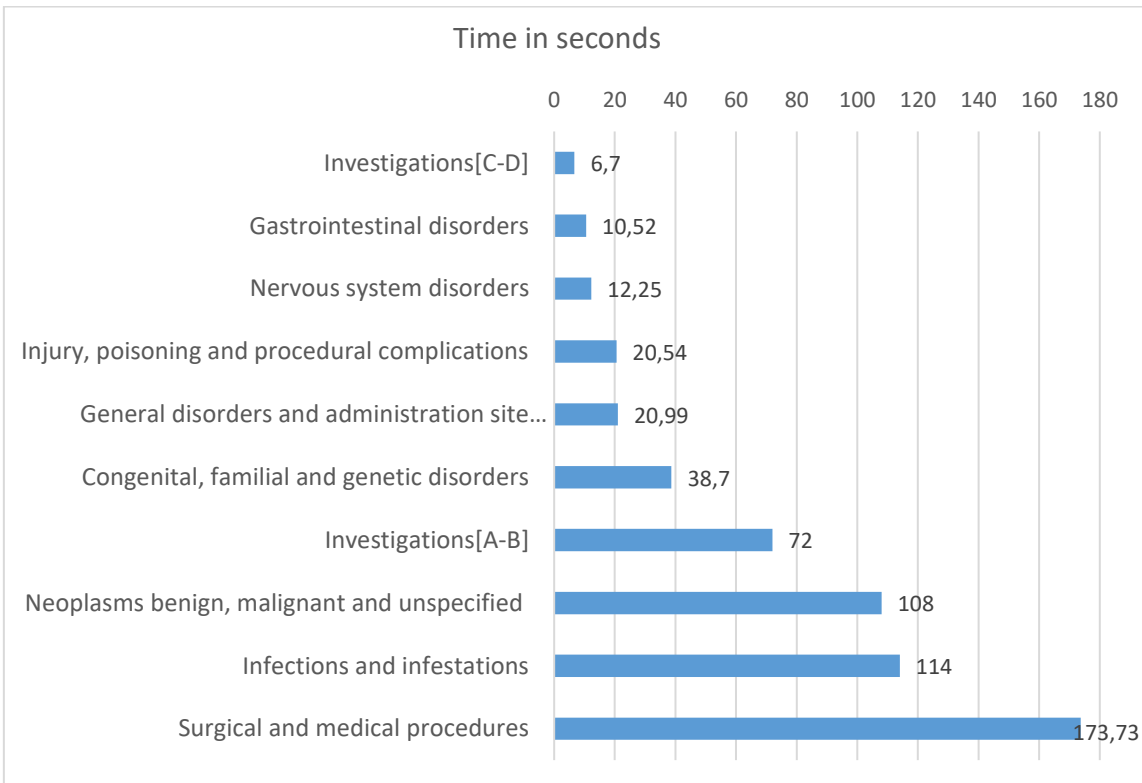


Figure 10 Loading times for reaction's groups

For the results in error, given the difficulty of assessing all the records one at the time, a set of previously seen or drugs that, through the shared experience of the Porto Pharmacovigilance unit and researchers, could be good targets for error evaluation. That said, 2 drugs classification was evaluated and the results are the one below.

1. **Paracetamol** this drug was tested, resulting in not finding drugs on several sites, while the test was done manually, results appeared for every site. The issue was that paracetamol was being searched as acetaminophen (mainly USA nomenclature) and there were sites where it was called paracetamol.
2. **Vaccines**, being that empiric reports during the interviews and the thesis process, vaccines had a great error margin. That said, evaluated the database, the Vaccines, of a total of 56 existent in Portuguese, only 15 had synonyms and these did not return good matches. The reason is that products with no correspondence with English and Portuguese terms, are searched with the Portuguese name, generating a lot of error.

As for reproducibility, the system has a few issues. When it did show a result, it was the same every time. However, results that previously appeared, can lead into an error or timeout in a sequential search. This can be explained that the websites themselves can impact the system in more ways that predicted.

# 5. Analysis and Discussion

## 5.1. Future developments and insights

As stated earlier, all the efforts invested in this research are not only aimed at the current phase, but also to suggest a roadmap for the development of the tool. For this, all the feedback, impressions and experience collected in the interviews was added to my own experience and the state of the art. This was done taken into account the technologies already in use, the initial focus of the tool and the best practices of software development and pharmacovigilance.

One of the possibilities that could come in handy is the ability to create RESTful operations on top of the mongoDB. The notion of Representational State Transfer (REST) was first defined by Roy Fielding in 2000 as a set of design constraints, methods, and architectures that leads to scalable, reliable, easy to use interfaces (85). Since then, a community has developed that adopts many of the base principles of REST. These principles are (86):

- Uniform Interface: Individual resources are identified using URLs, and can be represented in multiple different ways like JSON or Extensible Markup Language (XML). Clients manipulate the resource through the representations using self-descriptive messages. Hypermedia (hyperlinks) and hypertext act as the engine for state transfer.
- Stateless Interactions: None of the client's context is stored on the server side between requests, so all of the information necessary to service the request is contained in the URL, headers, or body.
- Cacheable: Responses can be cached and responses must define themselves as cacheable or not.
- Client and Server are separated from each other so the client is not worried with the data storage while the server is not disturbed with the user interface.
- Layered System: At any time, it should not be possible to tell if one is connected to the end server or to an intermediate. Intermediaries can help enforce the security policies, enable load-balancing, etc.

Implementing this architecture for the data collected could transform the system into a universal API for substitution of the regular sites. This could create the beginning of a paradigm shift, creating accessible reporting data to the public (when possible, due to identification restrictions). The implementation could be handled along with the industry standards, creating a data model for incorporating the Health Level 7 (HL7) Fast Healthcare Interoperability Resources (FHIR) and/or International Organization for Standardization (ISO) identification of Medicinal Products (IDMP) models into the API.

FHIR is a specification based on the RESTful principles and is organized around the concept of "Resources" and defines many types of resources that describe the healthcare space. In the FHIR standard, all resources have references to other resources, extensions (that can be made by everyone) and human readable texts. FHIR has an open license, a focus on implementation, and a formal maturity process linked to implementation outcome (86). It is intended to be the next healthcare standard for communication and its adoption has been spreading hugely in the last few years, with companies like google and apple using it (87).

IDMP standards are intended for standardize definitions for the identification and description of medicinal products. Their purpose is to help the reliable exchange of

medicinal product information in a robust and consistent manner. This reliability in the exchange ensures wide interoperability across worldwide regulatory and healthcare communities, which is critical in ensuring accurate analysis and unambiguous communication across jurisdictions (88).

This format of communication along the standard implementation would enhance the capability of communicating easily with healthcare biomedical systems and facilitate interaction since the standards are free to use and heavily documented by a thriving community. Moreover, such an open approach and ease of use related to the data, would encourage everyone worldwide to analyse the ADR reports and enable the creation new knowledge based on it.

Along the lines of the RESTful platform, it could also be the corner stone of multiple services of the system on top of it, like a chat bot. A chatbot is a service, powered by rules and sometimes artificial intelligence, that you interact with via a chat interface. This service is a computer program designed to interact with people by closely emulating human conversation, as idealized in the Turing Test for AI and it could live in any major chat product (Facebook Messenger, Slack, Text Messages, etc.). The interest in chatbots has been increasing recently (89) and it is now very common to see notification bubbles pop on shopping. This is no surprise since the latest reports state that these kinds of bot achieve conversion rates higher than any other system tested at a fraction of the cost (90). This could be a very handy tool to increase reporting “conversion rates”. A chat bot could use the information stored inside the system to answer small questions for the users of the regional offices website, social media pages or even text messages. Actions like small troubleshooting and Frequent Asked Questions (FAQ) could be a way of increasing the connection with the community and close the gap between the audience and the officers. The use of chatbot in healthcare settings is not an unprecedented application. There are reports of chatbot used with reasonable success on accompanying cancer patients (91) and for supporting psychiatric patients (92). Another feasible improvement for the current system is connected with the information present in the system, which could be augmented and provide more information to the officers.

One of the sources already tried is the biomedical literature – a source that regional offices often do not have the tools to gather. The process has been implemented in some cases studies, mainly recurring to text-mining techniques. This technique is defined as the process of extracting meaningful information from large amounts of unstructured text using computational methods (93). Because text mining provides a mechanism to transform free-text into computable knowledge, text mining is emerging as a way to explore, analyse, query, and manage underutilized safety information about drugs (94) and ADR. To meet the challenges posed by unstructured text, text mining employs a wide range of statistical, machine learning, and linguistic techniques that are associated with natural language processing (NLP). It is beneficial to think of text mining as a process that uses tools, methods, and heuristics developed by those who research the processing of natural language.

One argument in favour of these applications is the language barrier. Since most biomedical literature is written in English, most algorithms are developed for it as well, facilitating the access to worldwide algorithms with ease. One argument against it could be the usage intended when the algorithms are developed, like different databases, or high processing requirements. One of the best cases found is related to a probabilistic

method that given a drug-effect pair searches PubMed for abstracts and converts these abstracts to standard NLP features (46).

This is a feasible add-on to the system, since the database is very change-tolerant and the retrieving process can be changes easily as well, being that a major issue could the times associated with the searches or search method. Instead of a point-and-click issue, it could be a gradual search for a set and trying to enlarge the database as time went by, in order to suggest papers with high correlation rate with the search made. This could prevent a great loading time for the “real-time” search that is implemented in the moment, that already takes a fairly large amount of time.

Another source is social media that, as stated, already has been a target for ADE monitoring. The main focus has been Twitter, since it is dedicated to a real-time share of thoughts and opinions, high volume of users and data publicly available to third parties. Reports suggest a relation between ADR described on twitter and the ones reported to the Pharmacovigilance system of the USA (95). At least 4 use cases have been described that correlates “twits” with a drug and an adverse reaction (96,97). Nevertheless, other sources of social-media have been described, such as internet-forums, blogs and health websites (98). These algorithms already bring a little more implementation system, since the language techniques were not built for Portuguese language and this can bring a large variation into the results, since the screened target population would not be Portuguese. Nevertheless, if properly labelled, this could be useful information for pharmacovigilance officers to assess with the due caution.

In sum, the versatility of the system enables a large improvement, providing a framework to improve its capabilities. The proposed additions not only deepen the information provided in terms of quality and sources, but also increase the actual scope of the tool.

## 5.2. Interviews

An overall message perceived from the interviews is that healthcare professionals understand the importance of notification, but is still not something that they deal on a daily basis. Another generic feeling throughout the conversation is that a closer connection with the regional centres could be key for improving these numbers. Furthermore, there are health institutions that are clearly winning cases and should be looked upon more closely. The main focuses talked along the session were notification number and how to increase them, the roadmap proposed and if technology could help pharmacovigilance.

### 5.2.1. Notification of ADR

Most healthcare professionals reported that ADR notification was not a systematic practice in their day-to-day activities. For the physician, information regarding the possible ways of reporting was not very clear and feels his colleagues felt the same way. The nurse felt the same way, expressing that in her activities, it was very difficult to assess a possible ADR in her patients. The opinion of the developer, the healthcare professionals were not evangelized enough, reporting that the times working with an ADR reporting tool, he noticed healthcare professionals were drastically more report-active if someone in their workgroup acted as a promoter for the notification.

Additionally, a fact stated by more than 1 healthcare professional was that information regarding notification was scarce, and especially the physician felt that most

of the information that he received was coming from the pharmaceutical industry and not from the governmental entities.

As for the hospital pharmacist, belonged to a very active hospital, being that she felt that her and her colleagues were so active due to a close relation with the regional office, a very active promotion of the pharmacovigilance in the pharmacy service and the system that they had access to that enabled a fast reporting (SIRAI). Additionally, she stated that a clinical pharmacist's role, like the one the pharmacists had in her service was also a factor for increasing notification. The relation with the physician and remaining healthcare team was important for improving awareness for notification, quality of the notification and the numbers themselves.

The community pharmacist had not such a way of reporting and the main reason for not being so active in the ADR reporting was time. The fact that always being in the balcony was a major issue regarding reporting behaviour. Furthermore, when asked if a system could be helpful for improving reporting, the community pharmacist said definitely and a link of integration with the pharmacy software would be a major improvement.

#### 5.2.2. Information Systems for helping ADR notification

The main focus for an information system in the opinion of all the inquiries was speed. Loading times and a lot of clicks was considered a main problem regarding health information system and a system designed for promoting notifications should have those metrics as quartertones. Regarding the current causality assessment tool, most healthcare professionals felt it should exist in order to facilitate the work of the regional offices, but the usefulness to them was not yet clear or neglected. One major issue and of general appraisal was that a tool that would be fast and of very simple usage would benefit the notification number dramatically. The physician gave the example of the tool for mandatory infections reporting, the "Sistema Nacional de Vigilância Epidemiológica" (SINAVE) that was a very good tool and a possible example to follow in order to increase notifications. The issue regarding the SINAVE was very curious since it is a similar concept of the ADR notification. Even though clinicians are bound by law to do so, the need to create an efficient method for notification is indeed similar. SINAVE was created by the Portuguese government to dematerialize the mandatory notification of transmissible diseases and other public health risks and enables physician to report in real time the occurrence of the disease so the appropriate measure could be taken. Furthermore, it acts as an instrument of continuous monitorization for these illnesses in the whole country (99). The fact that is implemented in the prescription software, helps the create awareness, being that is available a weblink to it, fastening the access and availability. Images of SINAVE are available in the ANNEX B

The physician reported as well that nowadays there are a few mobile apps that can be easily accessed in order to retrieve information or take specific actions. He used it as an example as well.

The hospital pharmacist stated that a system for creating notification greatest features should be reliability and speed. Interesting features would be the capability of seeing the history of reports and the result of the causality assessment. Moreover, being able to check metrics of the reports would be a plus. The hospital pharmacist indicated that being trustworthy is the single most interesting feature, but usability, simplicity of use is second to it. But the road to follow according to her is to integrate these kinds of

tools into the information system to make notifications more automatic, or at least having pre-filled fields.

The pharmacovigilance officer, which was the only one with access and experience with the tool, referred it was a very useful tool to her job but it was very slow at the moment. Furthermore, vaccines were a problem to address. Nevertheless, she seemed excited with the possibilities this tool could offer. As for additional features, she felt that helping with the creation of feedback for the reporters could be a good option to follow and adding the option of categorizing different reporters according to their professional status. Sending scientific literature for healthcare professionals could be very useful and sending the same articles for patients could cause confusion due to the fact that they could not be capable of interpreting the values given.

Overall, the general assumption perceived felt like notification system and this system would be perfect match, creating a window of opportunity for health information system, or regulatory entities to create a macro pharmacovigilance tool to embrace all of the notification workflow.

### 5.2.3. Roadmap

As for the proposed roadmap and functionalities, the most active one was the developer, who clearly was on a more familiar setting. The global response to the proposed features was enthusiastic, but with a grain of caution. He advised that such features, especially the one requiring machine-learning and statistic-heavy features would not be simple and were very ambitious, even though he felt it would be worth it. The API proposal was clearly the high point, since he felt that “sharing was caring” and in his opinion the advantages clearly outweighed the disadvantages. The facilitation of access to information was crucial and he felt, that clinicians and healthcare professionals in Portugal were starting to grasp the benefits of big data and analytics on top of it. That was clearly a way of motivating people to notify, a method of showing a return of investment, creating value for the pharmacovigilance offices, but also for the notifier. So not only feedback could be useful, but also metrics could become a way of providing value for the notifier.

One major warning that was left in the conversation with the developer was that hardware, at least in the public healthcare centres, was old and with very low specification, impacting deeply the performance and usability.

As for the pharmacovigilance officer, felt that adding biomedical literature could be very helpful, giving more information to the officer and in last instance to the reporter. As for chatbot, she felt it could have a negative impact, as it could create a sense of reporting while texting the chatbot, when it should be only a starting point. Nevertheless, she agreed that something that was not human-resource heavy and create a sense of awareness and proximity to the audience could be helpful.

### 5.3. Technical capabilities

The feature evaluation showed very distinct behaviours for the websites. MEDSCAPE is not impacted by group size, since its search does not contemplate groups or even MedDRA coding and simply searches for the drug on the website. The search is direct and its low loading time is a clear proof of that.

As for the rest, more reactions impact the loading time substantially. Against what could be a first response, different groups seem to impact less than reactions of the

same group. Additionally, there are examples of the MHRA where 5 reactions from the same group had a loading time drastically inferior to the one with 5 reactions from different groups. As for vigiaccess, the loading times seem always similar, being that the differences can be explained by the polling mechanism of the front-end, currently set for 10 seconds. So, if a result was ready right after a polling, it will only appear in 10 seconds, beside already being ready.

As for errors, the EMA website is clearly affected by the current methodology. Since the website is flash based and loads the content on the fly, causes several problems on the rendering system of the webdriver, recurrently not finding the correct box to tick or graph to click. Vigiaccess also suffers from error returning, but this issue can be solved since an API has been developed by the WHO and can be accessed by the certified regional pharmacovigilance offices. This can facilitate the access to the data and should be clearly in the development roadmap for the short-term.

Regarding the capability of not returning results, the most impacted website is the MEDSCAPE. This can be explained by nomenclature issues, since the website is north-American. North-American institutions have different naming for drugs from the European entities. This can be resolved by a rearranging of the translation system and create more website specific nomenclatures besides Portuguese-English. MHRA is second on non-finding drugs, also due to a naming issue. The format of names on the website are somewhat different, non-including salt formation and compounds. This also can be solved by a stricter dictionary of synonyms.

Nevertheless, it should be assessed with the Porto Pharmacovigilance unit if more searches can be made in order to better evaluates these impacts and finding uncorrelated synonyms.

## 6. Conclusions

As the thesis went on, it became clear that pharmacovigilance will gain a great amount by pairing with technology. With the normalization of healthcare application in the healthcare institutions, but more than that, in the personal space, it becomes a necessity to bind pharmacovigilance to these developments. The healthcare professionals feel the same way and the global trends indicate a similar path. The issue now is to gather the consensus of the legislators and decision-makers to commit to it.

The current system is seen as a very useful piece of technology that could make an enormous impact into the NPS, not only in the Porto regional office. It already provides an easy way of gathering information but it still lacks some improvement in order to have a total adoption rate. The loading time is not yet reasonable in some cases and the ratio of error returning makes the tool unreliable to using. But as these issues are addressed, the system can be improved upon and this is where it really could shine. Creating a framework for open data, or adding more information of distinct sources is a real possibility and studies show it can be done. Taping into English written artefacts is in this case a big plus, because partnerships and synergies can be more easily established around the world. As for the rest, a chatbot could be interesting but more research is needed. As for the API, it is a certainty that sharing is caring, and involving more people into this cause, can only be helpful. Finally, a framework for further developing the tool could be helpful, for creating a common ground of development and stabilising the source code. Through this method, logging the metrics discussed in here could be essential for assessing real improvements of the tool.

As for notification numbers, it became clear along this thesis that an ease-of-use system integrated into the health information systems, whether hospital pharmacy, community pharmacy, clinical services or even a mobile app, could benefit greatly the notification numbers. In this day and age, it becomes difficult for the users to grasp the need for repeating the input of information and as time passes by, it becomes a greater barrier for adoption of technology in the healthcare setting.

## 7. References

1. UMC | Common concepts and terms [Internet]. [cited 2018 May 25]. Available from: <https://www.who-umc.org/safer-use-of-medicines/safer-use-of-medicines-the-basics/common-concepts-and-terms/>
2. Nebeker JR, Yarnold PR, Soltysik RC, Sauer BC, Sims SA, Samore MH, et al. Developing Indicators of Inpatient Adverse Drug Events Through Nonlinear Analysis Using Administrative Data. *Med Care* [Internet]. 2007;45(Suppl 2):S81–8. Available from: <http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=0005650-200710002-00015>
3. Formica D, Sultana J, Cutroneo PM, Lucchesi S, Angelica R, Crisafulli S, et al. The economic burden of preventable adverse drug reactions: a systematic review of observational studies [Internet]. Vol. 17, Expert Opinion on Drug Safety. Taylor & Francis; 2018. 681–695 p. Available from: <https://doi.org/10.1080/14740338.2018.1491547>
4. Chiatti C, Bustacchini S, Furneri G, Mantovani L, Cristiani M, Misuraca C, et al. The economic burden of inappropriate drug prescribing, lack of adherence and compliance, adverse drug events in older people a systematic review. *Drug Saf.* 2012;35(SUPPL. 1):73–87.
5. Shehab N, Lovegrove MC, Geller AI, Rose KO, Weidle NJ, Budnitz DS. US emergency department visits for outpatient adverse drug events, 2013–2014. *JAMA - J Am Med Assoc.* 2016;316(20):2115–25.
6. Ritter J, Lewis L, Mant T, Ferro A. A textbook of clinical pharmacology and therapeutics. CRC Press; 2008.
7. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: A meta- analysis of prospective studies. *J Am Med Assoc.* 1998;279(15):1200–5.
8. Angamo MT, Chalmers L, Curtain CM, Bereznicki LRE. Adverse-Drug-Reaction-Related Hospitalisations in Developed and Developing Countries: A Review of Prevalence and Contributing Factors. *Drug Saf.* 2016;39(9):847–57.
9. Scripcaru G, Mateus C, Nunes C. A decade of adverse drug events in Portuguese hospitals: Space-time clustering and spatial variation in temporal trends. *BMC Pharmacol Toxicol.* 2017;18(1):1–8.
10. Baldini A, Von Korff M, Lin EHB. A Review of Potential Adverse Effects of Long-Term Opioid Therapy. *Prim Care Companion CNS Disord* [Internet]. 2012 Jun 14; Available from: <http://www.psychiatrist.com/pcc/article/pages/2012/v14n03/11m01326.aspx>
11. Manchikanti L, Abdi S, Atluri S, Balog CC, Benyamin RM, Boswell M V, et al. American Society of Interventional Pain Physicians (ASIPP) guidelines for responsible opioid prescribing in chronic non-cancer pain: Part I--evidence assessment. *Pain Physician* [Internet]. 2012;15(3 Suppl):S1-65. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22786448>
12. Kahan M, Mailis-Gagnon A, Wilson L, Srivastava A. Canadian guideline for safe and effective use of opioids for chronic noncancer pain: Clinical summary for family physicians. *Can Fam Physician.* 2011;5757:1257–66.

13. Frontier Economics. Exploring the costs of unsafe care in the NHS. *Front Econ London* [Internet]. 2014;(October):1–31. Available from: <http://www.frontier-economics.com/documents/2014/10/exploring-the-costs-of-unsafe-care-in-the-nhs-frontier-report-2-2-2-2.pdf>
14. European Commission. Proposal for a regulation amending, as regards pharmacovigilance of medicinal products for human use. Regulation (EC) No 726/2004. Impact assessment. 2008;l(December):12. Available from: [https://ec.europa.eu/health/sites/health/files/files/pharmacos/pharmpack\\_12\\_2008/pharmacovigilance-ia-vol1\\_en.pdf](https://ec.europa.eu/health/sites/health/files/files/pharmacos/pharmpack_12_2008/pharmacovigilance-ia-vol1_en.pdf)
15. Bouvy JC, De Bruin ML, Koopmanschap MA. Epidemiology of Adverse Drug Reactions in Europe: A Review of Recent Observational Studies. *Drug Saf* [Internet]. 2015;38(5):437–53. Available from: <http://dx.doi.org/10.1007/s40264-015-0281-0>
16. Poleksic A, Xie L. Predicting serious rare adverse reactions of novel chemicals. *Bioinformatics* [Internet]. 2018 Mar 29;14(3):10.4088/PCC.11m01326. Epub 2012 Jun 14. Available from: <https://academic.oup.com/bioinformatics/advance-article/doi/10.1093/bioinformatics/bty193/4956346>
17. Vargesson N. Thalidomide-induced teratogenesis: History and mechanisms. *Birth Defects Res Part C - Embryo Today Rev.* 2015;105(2):140–56.
18. McBride W. Thalidomide and congenital abnormalities. *Lancet.* 1961;2(3):1358–1358.
19. Van Grootheest K. The dawn of pharmacovigilance: An historical Perspective. *Int J Pharm Med.* 2003;17(5–6):195–200.
20. WHO. WHO-UMC | Glossary [Internet]. [cited 2018 May 25]. Available from: <https://www.who-umc.org/global-pharmacovigilance/global-pharmacovigilance/glossary/>
21. Harpaz R, DuMochel W, Shah NH. Big data and adverse drug reaction detection. *Clin Pharmacol Ther.* 2016;99(3):268–70.
22. Harpaz R, DuMouchel W, Shah NH, Madigan D, Ryan P, Friedman C. Novel Data-Mining Methodologies for Adverse Drug Event Discovery and Analysis. *Clin Pharmacol Ther* [Internet]. 2012 Jun 2;91(6):1010–21. Available from: <http://doi.wiley.com/10.1038/clpt.2012.50>
23. Duarte M, Ferreira P, Soares M, Cavaco A, Martins AP. Community pharmacists' attitudes towards adverse drug reaction reporting and their knowledge of the new pharmacovigilance legislation in the southern region of Portugal: a mixed methods study. *Drugs Ther Perspect.* 2015;31(9):316–22.
24. Corrêa-Nunes AM. O sistema de farmacovigilância em Portugal (sua criação e desenvolvimento). *Cad Saude Publica.* 2005;14(4):725–33.
25. Hazell L, Shakir SAW. Under-reporting of adverse drug reactions : a systematic review. *Drug Saf* [Internet]. 2006;29(5):385–96. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16689555>
26. Ribeiro-vaz I. Using Information Systems in Pharmacovigilance Promotion of Adverse Drug Reaction Reporting. University of Porto; 2016.
27. Ribeiro-Vaz I, Silva AM, Costa Santos C, Cruz-Correia R. How to promote adverse drug reaction reports using information systems - A systematic review and meta-analysis. *BMC Med Inform Decis Mak* [Internet]. 2016;16(1):1–10.

- Available from: <http://dx.doi.org/10.1186/s12911-016-0265-8>
28. Infarmed IP. Evolução das notificações de RAM no SNF por origem [Internet]. [cited 2019 Apr 25]. Available from: <http://www.infarmed.pt/web/infarmed/entidades/medicamentos-uso-humano/farmacovigilancia/desempenho-do-snf>
  29. Infarmed IP. Envio de reações adversas para o Eudravigilance - Novas funcionalidades [Internet]. 2017 [cited 2019 Jul 5]. p. 1. Available from: <http://www.infarmed.pt/documents/15786/1539452/Envio+de+reações+adversas+para+o+Eudravigilance+-+Novas+funcionalidades/f8a4c5ff-a48e-4b7c-ab1c-4743b89e90f0>
  30. Toki T, Ono S. Spontaneous Reporting on Adverse Events by Consumers in the United States: An Analysis of the Food and Drug Administration Adverse Event Reporting System Database. *Drugs - Real World Outcomes* [Internet]. 2018;5(2):117–28. Available from: <https://doi.org/10.1007/s40801-018-0134-0>
  31. Banovac M, Candore G, Slattey J, Arlett P, Houžez F, Haerry D GG. Patient reporting in EudraVigilance – a measure of patient engagement? Work Meas impact Pharmacovigil Act [Internet]. 2016;(December). Available from: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Presentation/2016/12/WC500218541.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Presentation/2016/12/WC500218541.pdf)
  32. Centrelareb netherlands pharmacovigilance. Highlights 2016. 2016;1–12.
  33. Infarmed IP et al. Farmacovigilância em Portugal: 25 anos. 2018.
  34. Ribeiro-Vaz I, da Silva F, Silva A-M, Alves D, Cruz-Correia R. Pharmacovigilance Informatics. In 2016. p. 300–16.
  35. Beninger P, Ibara MA. Pharmacovigilance and Biomedical Informatics: A Model for Future Development. *Clin Ther* [Internet]. 2016 Dec;38(12):2514–25. Available from: <http://www.sciencedirect.com/science/article/pii/S014929181630844X>
  36. NET MARKETSHARE. Browser Market Share [Internet]. [cited 2019 Aug 25]. Available from: <https://www.netmarketshare.com/browser-market-share.aspx?options=%7B%22filter%22%3A%7B%22%24and%22%3A%5B%7B%22deviceType%22%3A%7B%22%24in%22%3A%5B%22Desktop%2Flaptop%22%5D%7D%7D%5D%7D%2C%22dateLabel%22%3A%22Custom%22%2C%22attribute%22%3A%22share%22%2C%22>
  37. FDA. FDA Adverse Events Reporting System Electronic Submissions. [Internet]. [cited 2019 Aug 25]. Available from: <https://www.fda.gov/drugs/questions-and-answers-fdas-adverse-event-reporting-system-faers/fda-adverse-event-reporting-system-faers-electronic-submissions>
  38. EMA. ICH guideline E2B (R3) on Electronic Transmission of Individual Case Safety Reports (ICSRs) - data elements and message specification - implementation guide [Internet]. [cited 2019 Aug 25]. Available from: [https://www.ema.europa.eu/en/documents/scientific-guideline/international-conference-harmonisation-technical-requirements-registration-pharmaceuticals-human-use\\_en-4.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/international-conference-harmonisation-technical-requirements-registration-pharmaceuticals-human-use_en-4.pdf)
  39. Innovative Medicine Initiative [Internet]. [cited 2019 Aug 25]. Available from: <https://www.imi.europa.eu/>
  40. Dal Pan GJ. Ongoing challenges in pharmacovigilance. *Drug Saf*. 2014;37(1):1–8.
  41. Li Y, Ryan PB, Wei Y, Friedman C. A Method to Combine Signals from

- Spontaneous Reporting Systems and Observational Healthcare Data to Detect Adverse Drug Reactions. *Drug Saf.* 2015;38(10):895–908.
42. Schmider J, Kumar K, LaForest C, Swankoski B, Naim K, Caubel PM. Innovation in Pharmacovigilance: Use of Artificial Intelligence in Adverse Event Case Processing. *Clin Pharmacol Ther.* 2018;105(4).
  43. Bollegala D, Maskell S, Sloane R, Hajne J, Pirmohamed M. Causality Patterns for Detecting Adverse Drug Reactions From Social Media: Text Mining Approach. *JMIR public Heal Surveill* [Internet]. 2018 May 9 [cited 2018 Aug 10];4(2):e51. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29743155>
  44. Yang M, Kiang M, Shang W. Filtering big data from social media - Building an early warning system for adverse drug reactions. *J Biomed Inform* [Internet]. 2015;54:230–40. Available from: <http://dx.doi.org/10.1016/j.jbi.2015.01.011>
  45. Luo Y, Thompson WK, Herr TM, Zeng Z, Berendsen MA, Jonnalagadda SR, et al. Natural Language Processing for EHR-Based Pharmacovigilance: A Structured Review. *Drug Saf.* 2017;40(11):1075–89.
  46. Natarajan S, Bangera V, Khot T, Picado J, Wazalwar A, Costa VS, et al. Markov logic networks for adverse drug event extraction from text. *Knowl Inf Syst* [Internet]. 2017 May 8;51(2):435–57. Available from: <http://link.springer.com/10.1007/s10115-016-0980-6>
  47. Segura-Bedmar I, Martínez P. Pharmacovigilance through the development of text mining and natural language processing techniques. *J Biomed Inform.* 2015;58:288–91.
  48. Shang N, Xu H, Rindflesch TC, Cohen T. Identifying plausible adverse drug reactions using knowledge extracted from the literature. *J Biomed Inform* [Internet]. 2014;52:293–310. Available from: <http://dx.doi.org/10.1016/j.jbi.2014.07.011>
  49. Drosatos G, Kaldoudi E. Blockchain Applications in the Biomedical Domain: A Scoping Review. *Comput Struct Biotechnol J* [Internet]. 2019;17:229–40. Available from: <https://doi.org/10.1016/j.csbj.2019.01.010>
  50. Min J, Osborne V, Lynn E, Shakir SAW. First Conference on Big Data for Pharmacovigilance. *Drug Saf* [Internet]. 2018;41(12):1281–4. Available from: <https://doi.org/10.1007/s40264-018-0727-2>
  51. Bernstam E V., Smith JW, Johnson TR. What is biomedical informatics? *J Biomed Inform* [Internet]. 2010;43(1):104–10. Available from: <http://dx.doi.org/10.1016/j.jbi.2009.08.006>
  52. MedTech Europe. The European Medical Technology industry - in figures 2019. *MedTech Eur - from diagnosis to cure* [Internet]. 2019;44. Available from: <https://www.medtecheurope.org/wp-content/uploads/2019/04/The-European-Medical-Technology-Industry-in-figures-2019-2.pdf>
  53. Accenture. Accenture 2018 Consumer Survey on Digital Health [Internet]. 2019 [cited 2019 Apr 25]. Available from: [https://www.accenture.com/us-en/insight-new-2018-consumer-survey-digital-health?utm\\_source=newsletter&utm\\_medium=email&utm\\_campaign=newsletter\\_axiosvitals&stream=top-stories](https://www.accenture.com/us-en/insight-new-2018-consumer-survey-digital-health?utm_source=newsletter&utm_medium=email&utm_campaign=newsletter_axiosvitals&stream=top-stories)
  54. App Annie [Internet]. [cited 2019 Jul 25]. Available from: <https://www.appannie.com/en/insights/market-data/medical-apps-transform-patient-care/>

55. IQVIA. The Growing Value of Digital Health [Internet]. [cited 2019 Apr 25]. Available from: <https://www.iqvia.com/institute/reports/the-growing-value-of-digital-health>
56. Shuren J, Patel B, Gottlieb S. FDA regulation of mobile medical apps. *JAMA - J Am Med Assoc.* 2018;320(4):337–8.
57. Administration F and D. Guidance for Industry and Food and Drug Administration Staff Preface Public Comment. 2016; Available from: <http://www.regulations.gov>.
58. Medicines & Healthcare Products Regulatory Agency. Guidance: Medical device stand-alone software including apps. 2018;1–39. Available from: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/717865/Software\\_flow\\_chart\\_Ed\\_1-05.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/717865/Software_flow_chart_Ed_1-05.pdf)
59. Ovretveit J. Medical managers can make research-based management decisions. *J Manag Med.* 1998;12(6):391–7, 322.
60. Ammenwerth E, Brender J, Nykänen P, Prokosch HU, Rigby M, Talmon J. Visions and strategies to improve evaluation of health information systems: Reflections and lessons based on the HIS-EVAL workshop in Innsbruck. *Int J Med Inform.* 2004;73(6):479–91.
61. Han YY. Unexpected Increased Mortality After Implementation of a Commercially Sold Computerized Physician Order Entry System. *Pediatrics* [Internet]. 2005;116(6):1506–12. Available from: <http://pediatrics.aappublications.org/cgi/doi/10.1542/peds.2005-1287>
62. Friedberg MW, Chen PG, Van Busum KR, Aunon F, Pham C, Caloyeras J, et al. Factors Affecting Physician Professional Satisfaction and Their Implications for Patient Care, Health Systems, and Health Policy [Internet]. RAND Corporation. 2013. 1–152 p. Available from: [http://www.rand.org/pubs/research\\_reports/RR439.html](http://www.rand.org/pubs/research_reports/RR439.html)
63. Campbell E, Sittig DF, Ash JS, Guappone KP, Dykstra RH. Types of Unintended Consequences Related to Computerized Provider Order Entry. *Jamia.* 2006;13(5):547–56.
64. Friedman CP, Wyatt JC. *Evaluation Methods in Biomedical Informatics (Health Informatics)*. Berlin, Heidelberg: Springer-Verlag; 2005.
65. Cimino JJ, Shortliffe EH. *Biomedical Informatics: Computer Applications in Health Care and Biomedicine (Health Informatics)*. Berlin, Heidelberg, Heidelberg: Springer-Verlag; 2006.
66. Buckley JW, Buckley MH, Chiang H-F, Accountants. NA of, Canada. S of IA of. *Research methodology & business decisions*. [Canada]: National Association of Accountants; 1976.
67. Crotty M. *The foundations of social research : meaning and perspective in the research process*. London; Thousand Oaks, Calif.: Sage Publications; 1998.
68. Noor KBM. Case Study: A Strategic Research Methodology. *Am J Appl Sci* [Internet]. 2008 Nov 1;5(11):1602–4. Available from: <http://www.thescipub.com/abstract/?doi=ajassp.2008.1602.1604>
69. House ER. *Evaluating with validity* [Internet]. Sage Publications; 1980. Available from: <https://books.google.pt/books?id=HQIHAAAAMAAJ>
70. Shields P, Rangarjan N. *A Playbook for Research Methods: Integrating Conceptual Frameworks and Project Management*. Stillwater, OK: New Forums;

- 2013.
71. Uruno JONPF, Hu JIZ, Etersson DANEP, Inkelstein JOF. The Use and Interpretation of Quasi-Experimental Studies in Medical Informatics. 2006;16–23.
  72. Approach AE, Lau F, Kuziemyk C. Handbook of eHealth Evaluation: An Evidence-based Approach. 2016.
  73. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* [Internet]. 2008 Dec 30;337(nov11 2):a2390–a2390. Available from: <http://www.bmj.com/cgi/doi/10.1136/bmj.a2390>
  74. Strobe E, Vandembroucke JP, Elm E Von, Altman DG, Gøtzsche PC, Mulrow CD, et al. Strengthening the Reporting of Observational Studies in Explanation and Elaboration. 2007;18(6):805–35.
  75. Alasuutari P. The rise and relevance of qualitative research. *Int J Soc Res Methodol*. 2010;13(2):139–55.
  76. Fanning E. Formatting a Paper-based Survey Questionnaire : Best Practices. 2005;10(12).
  77. Johnson RB, Onwuegbuzie AJ, Turner LA. Toward a Definition of Mixed Methods Research. *J Mix Methods Res* [Internet]. 2007 Apr 29;1(2):112–33. Available from: <http://sk.sagepub.com/reference/research/n269.xml>
  78. Best Practices for Mixed Methods Research in the Health Sciences. 2018;1–61.
  79. Jamshed S. Qualitative research method - interviewing and observation. 2014;5(4):4–5.
  80. Ranney ML, Meisel ZF, Choo EK, Garro AC, Sasson C, Morrow Guthrie K. Interview-based Qualitative Research in Emergency Care Part II: Data Collection, Analysis and Results Reporting. Newgard C, editor. *Acad Emerg Med* [Internet]. 2015 Sep;22(9):1103–12. Available from: <http://doi.wiley.com/10.1111/acem.12735>
  81. Croxson CH, Ashdown HF, Hobbs FR. GPs' perceptions of workload in England: a qualitative interview study. *Br J Gen Pract* [Internet]. 2017 Feb;67(655):e138–47. Available from: <http://bjgp.org/lookup/doi/10.3399/bjgp17X688849>
  82. DeJonckheere M, Vaughn LM. Semistructured interviewing in primary care research: A balance of relationship and rigour. *Fam Med Community Heal*. 2019;7(2):1–8.
  83. Infarmed IP. Notificações de RAM recebidas no SNF, por origem evolução 1992-2016 [Internet]. [cited 2019 Apr 13]. Available from: <http://www.infarmed.pt/documents/15786/2297404/Notificações+de+RAM+recebidas+no+SNF%2C+por+origem+evolução+1992-2016/5eb2f553-b29e-4359-a9dc-36f001793af9?t=1510755290896>
  84. Qu SQ, Dumay J. The qualitative research interview. *Qual Res Account Manag*. 2011;8(3):238–64.
  85. Fielding RT. Architectural Styles and the Design of Network-based Software Architectures. University of California, Irvine; 2000.
  86. Benson T, Grieve G. Principles of Health Interoperability [Internet]. Cham: Springer International Publishing; 2016. (Health Information Technology Standards). Available from: <http://link.springer.com/10.1007/978-3-319-30370-3>
  87. Google. FHIR on Google Cloud Platform [Internet]. [cited 2019 Sep 2]. Available

- from: <https://cloud.google.com/healthcare/docs/concepts/fhir>
88. EMA. Data on medicines (ISO IDMP standards): Overview [Internet]. [cited 2019 Aug 28]. Available from: <https://www.ema.europa.eu/en/human-regulatory/overview/data-medicines-iso-idmp-standards-overview>
  89. Zhang S, Dinan E, Urbanek J, Szlam A, Kiela D, Weston J. Personalizing dialogue agents: I have a dog, do you have pets too? ACL 2018 - 56th Annu Meet Assoc Comput Linguist Proc Conf (Long Pap. 2018;1:2204–13.
  90. Steinbach R. Examples of successful e-commerce shopping chatbots [Internet]. 2018 [cited 2019 Aug 28]. Available from: <https://www.snaps.io/blog/examples-successful-e-commerce-shopping-chatbots>
  91. Chaix B, Bibault JE, Pienkowski A, Delamon G, Guillemassé A, Nectoux P, et al. When chatbots meet patients: One-year prospective study of conversations between patients with breast cancer and a chatbot. *J Med Internet Res*. 2019;21(5):1–7.
  92. Ho A, Hancock J, Miner AS. Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *J Commun*. 2018;68(4):712–33.
  93. Singh MP. The Practical Handbook of Internet Computing [Internet]. 2004. 1144 p. Available from: <https://books.google.de/books?id=XKF2gg9Wq08C%5Cnhttps://books.google.com/books?id=XKF2gg9Wq08C&pgis=1>
  94. Harpaz R, Callahan A, Tamang S, Low Y, Odgers D, Finlayson S, et al. Text Mining for Adverse Drug Events: the Promise, Challenges, and State of the Art. *Drug Saf* [Internet]. 2014 Oct [cited 2018 Aug 10];37(10):777–90. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25151493>
  95. Freifeld CC, Brownstein JS, Menone CM, Bao W, Filice R, Kass-Hout T, et al. Digital drug safety surveillance: Monitoring pharmaceutical products in Twitter. *Drug Saf*. 2014;37(5):343–50.
  96. Sarker A, Ginn R, Nikfarjam A, O'Connor K, Smith K, Jayaraman S, et al. Utilizing social media data for pharmacovigilance: A review. *J Biomed Inform* [Internet]. 2015;54:202–12. Available from: <http://dx.doi.org/10.1016/j.jbi.2015.02.004>
  97. Cocos A, Fiks AG, Masino AJ. Deep learning for pharmacovigilance: recurrent neural network architectures for labeling adverse drug reactions in Twitter posts. *J Am Med Informatics Assoc* [Internet]. 2017 Jul 1;24(4):813–21. Available from: <https://academic.oup.com/jamia/article/26/6/580/5453996>
  98. Lardon JJ, Abdellaoui R, Bellet F, Asfari H, Souvignet J, Texier N, et al. Adverse drug reaction identification and extraction in social media: A scoping review. *J Med Internet Res*. 2015 Jul;17(7):1–25.
  99. Direção Geral de Saúde. SINAVE [Internet]. [cited 2019 Sep 4]. Available from: <https://www.dgs.pt/servicos-on-line1/sinave-sistema-nacional-de-vigilancia-epidemiologica.aspx>

# ANNEXES

## ANNEX A – Images of the causality assessment tool

**Farmacovigilância** HOME PESQUISA HISTÓRICO

### ELEMENTOS DE PESQUISA

MEDICAMENTO (PT) / PRINCÍPIO ATIVO (PT/EN)

Medicamento(PT) / Princípio Ativo(PT/EN) (\*) Medicamentos fora de comercialização

REAÇÃO ADVERSA A PESQUISAR

- Sistema de Órgãos - - Reação Adversa -

PESQUISAR LIMPAR

© CINTESIS : Centro de Investigação em Tecnologias e Serviços de Saúde

**Farmacovigilância** HOME RAM HISTÓRICO

### CONSULTA

Medicamento Capoten ( Captopril ) [ 25 mg ] - Gastrointestinal disorders - Abdominal pain

#### VIGIACCESS

Reação	Nº de Casos
Abdominal pain	346

#### MHRA-UK

Reação	Single		Multiple	
	Nº de Casos	Fatais	Nº de Casos	Fatais
Abdominal pain	79	0	3	0

#### EMA

Reação	Nº de Casos
Abdominal pain	1

### ELEMENTOS DE PESQUISA

MEDICAMENTO (PT) / PRINCÍPIO ATIVO (PT/EN)

(\*) Medicamentos fora de comercialização

### REAÇÃO ADVERSA A PESQUISAR

✖ Abdominal pain

**PESQUISAR**

LIMPAR

© CINTESIS : Centro de Investigação em Tecnologias e Serviços de Saúde

## REAÇÕES ADVERSAS A MEDICAMENTOS

**PESQUISAR**

© CINTESIS : Centro de Investigação em Tecnologias e Serviços de Saúde

# ANNEX B – Images of SINAVE

SINAVE - Sistema de Informação

The screenshot shows the 'Identificação do doente' (Patient Identification) form. At the top left is the DGS logo (Direção-Geral da Saúde) with the text 'desde 1899'. At the top right, it says 'Utilizador: Super User, Sessão expira em: 30 minutos' and has links for 'Log Off' and 'Renovar Sessão'. Below the header is a navigation bar with buttons for 'Início', 'Criar Notificação', 'Lista de Tarefas', 'Pesquisar', and 'Administração'. The main content area is titled 'Vigilância > Pesquisa Avançada' and contains the 'Identificação do doente' form. The form has two input fields: 'Número de utente no SNS' and 'Número de Identificação Civil'. At the bottom of the form are buttons for 'Registrar Novo Doente' and 'Pesquisar'. On the left sidebar, it says 'SIVDOT v 1.0.0 Cluster VM02'.

The screenshot shows the 'Registrar Novo Doente' (Register New Patient) form. It is divided into several sections: 'Dados pessoais', 'Dados sobre a morada', and 'Dados sobre cartões de identificação'. The 'Dados pessoais' section includes fields for 'Primeiro nome', 'Apelido', 'Número de identificação civil' (with a 'Desconhecido' checkbox), 'Data de Nascimento', 'Sexo', 'Doença', 'Telefone', 'Telemóvel', 'Correio Electrónico', and 'Nacionalidade'. The 'Dados sobre a morada' section includes 'País', 'Distrito', 'Concelho', 'Freguesia', and 'Morada'. The 'Dados sobre cartões de identificação' section has a field for 'Adicionar número de cartão de identificação'. At the bottom right are buttons for 'Iniciar Notificação' and 'Sair'. On the left sidebar, it says 'SIVDOT v 0.0.1 Cluster Cluster#04#010'.

The screenshot shows the main navigation bar of the SINAVE system. The 'Criar Notificação' button is highlighted with a red circle. The navigation bar also includes 'Início', 'Lista de Tarefas', 'Pesquisar', and 'Administração'. Below the navigation bar, the text 'Início' is visible, followed by a large empty box containing the text 'Bem-vindo ao SINAVE-SIVDoT'. On the left sidebar, it says 'SIVDOT v 1.0.0 Cluster VM02'. At the top right, it says 'Utilizador: Super User, Sessão expira em: 30 minutos' and has links for 'Log Off' and 'Renovar Sessão'.

ANNEX C – Pharmaceutical Forms without translation

**Name**

**Colírio, comprimido e solvente para solução**

**Comprimido e cápsula mole**

**Emplastro para teste cutâneo**

**Glóbulos**

**Granulado para solução oral ou retal**

**Implante, pó para suspensão**

**Pó e veículo para suspensão para uso intravesical**

**Pó para solução vaginal**

**Pó para suspensão para implantação**

**Solução para hemodiálise, hemodiafiltração e hemofiltração**

**Aditivo para banho**

**Cápsula vaginal**

**Chá medicinal**

**Chá medicinal instantâneo**

**Comprimido revestido + Comprimido**

**Concentrado para solução cutânea**

**Creme retal**

**Emplastro para prova de provocação**

**Emulsão e suspensão para emulsão injetável**

**Emulsão oral**

**Emulsão para pulverização auricular**

**Equivalente de tecido vivo**

**Gás medicinal criogénico**

**Gel dental**

**Gerador de radionuclidos**

**Gotas auriculares ou colírio, solução**

**Gotas auriculares, solução em recipiente unidose**

**Gotas auriculares, suspensão**

**Gotas orais, emulsão**

**Granulado para solução oral em saqueta**

**Grânulos**

**Liofilizado para solução para perfusão**

**Líquido para irrigação vesical**

**Matriz para Implantação**

**N.D.**

**Pasta dentífrica**

**Pasta gengival**

**Pasta oral**

**Pó e solução para solução injetável**

**Pó e solvente para cola para tecidos**

**Pó e solvente para solução cutânea**

**Pó e veículo para suspensão injetável em seringa pré-cheia**

**Pó para líquido para irrigação vesical**

**Pó para solução retal**

**Pó para suspensão injetável + Suspensão injetável**

**Pó para suspensão oral ou retal**

**Pó para suspensão oral ou retal em saqueta**

**Pó para suspensão retal**

**Pó, solvente e matriz para matriz para implantação**

**Pomada retal + Supositório**

**Solução para conservação de órgãos**

**Solução para hemodiálise**

**Solução para hemodiálise ou hemofiltração**

**Solução para hemofiltração**

**Solução para lavagem oftálmica**

**Solução para precursor radiofarmacêutico**

**Solução para pulverização auricular**

**Solução para uso intravesical**

**Solvente/Veículo para uso parentérico**

**Suspensão e granulado efervescente para suspensão oral**

**Suspensão injetável em caneta pré-cheia**

**Suspensão injetável em cartucho**

**Suspensão para implantação**

**Suspensão para pulverização cutânea**

**Tampão vaginal medicamentoso**

SEDE ADMINISTRATIVA

FACULDADE DE **MEDICINA**

FACULDADE DE **CIÊNCIAS**

