Phase IV clinical trials

Hadrys, Armin

UNIVERSITY OF ZAGREB
SCHOOL OF MEDICINE

Armin Hadrys

Phase IV Clinical Trials

GRADUATION PAPER

Zagreb, 2021

This graduate thesis was written at the Department of Internal Medicine, Unit of Clinical Pharmacology, University Hospital Centre Zagreb, and School of Medicine, Zagreb, Croatia, mentored by Prof. Dr. Sc. Robert Likić and submitted for evaluation in the academic year 2020/2021.

Mentor: Prof. Dr. Sc. Robert Likič

Table of Contents

1. Introduction ***1***

2. Overview of Pre-Marketing Phases ***2***

3. Shortcomings of Preclinical Trials for Detecting Real World Effectiveness ***3***

4. Post-Marketing Surveillance ***7***

4.1 Phase IV Clinical Trials **7**

4.2 Different Phase IV Study Designs **11**

4.2.1 Post-Marketing Surveillance Studies 12

4.2.2 Large Simple Trials 13

4.2.3 Randomized Clinical Trial 14

4.2.4 Physician Experience Studies (PES) 15

5. Discussion ***15***

6. Conclusion ***19***

7. Bibliography ***20***

8. Biography ***22***

Abstract

A phase IV clinical trial is an important step in the surveillance of a new drug after its approval for release on the market. It is an essential part of post-marketing research that focuses on real-world effectiveness and pharmacovigilance, thus not only continuing previous studies, but complementing them.

Phase IV clinical trials differ significantly in their study design, their requirements, and their scientific demand from previous study phases. Phases I to III mainly examine the safety profile of the drug on a smaller scale, as well as the drug's efficacy in a controlled environment of an RCT. Contrarily, phase IV studies claim to uncover even rarer ADRs that could have been overlooked in previous studies and also inspect whether the new drug proves its worth on the free market, in interaction with other drugs and in population groups that had not been previously admitted to the study.

The aim of this article is to set phase IV clinical trials in conjunction with the previous preclinical trials to get a bigger picture of the long path a drug takes—not only until its release on the market, but also after its commercialization. Other goals are to show the relation with different parts of post-marketing research, to investigate the modern role of phase IV studies, and to research the extent to which the current status of phase IV clinical trials meets requirements.

Keywords: phase IV clinical trials, post-marketing research, ADR, preclinical trials, pharmacovigilance, large simple trials, external validity

Abbreviations:

RCT – Randomized Controlled Trial
LST – Large Simple Trials
NDA – New Drug Application
FDA – Food and Drug Administration
EMA – European Medicines Agency
ADR – Adverse Drug Reaction
ADE – Adverse Drug Event

PSUR – Periodic Safety Update Report
PMS – Post-Marketing Surveillance Study

Sazetak

Klinička ispitivanja faze IV važan su korak u nadzoru novog lijeka, a nakon što je odobreno njegovo plasiranje na tržište.
Klinička ispitivanja faze IV bitan su dio istraživanja lijekova u njihovoj postmarketinškoj fazi (nakon dolaska lijeka na tržište), a usmjerena su na stvarnu učinkovitost i farmakovigilanciju, te stoga, ne samo da nastavljaju prethodna ispitivanja, već ih nadopunjuju.

Značajno se razlikuju u načinu ispitivanja, zahtjevima i znanstvenoj potražnji u odnosu na prethodna ispitivanja, koja su uglavnom ispitala sigurnosni profil lijeka u manjoj mjeri, kao i djelotvornost lijeka u kontroliranom okruženju RCT-a. S druge strane, ispitivanja faze IV, za koja se tvrdi da otkrivaju rijetke nuspojave, mogla bi biti zanemarena u prethodnim ispitivanjima i istraživanju na koji način novi lijek pokazuje svoju vrijednost na slobodnom tržištu u interakciji s drugim lijekovima i populacijskim skupinama, koje prethodno nisu bile uključene u ispitivanje.

Cilj rada je postaviti klinička ispitivanja faze IV u vezu zajedno s prethodnim pretkliničkim ispitivanjima, kako bi se dobila veća slika dugog puta kojim lijek prolazi - ne samo do trenutka njegova plasiranja na tržište - već i nakon komercijalizacije. Također, rad objašnjava odnos s drugim dijelovima postmarketinškog istraživanja, analizira ulogu ispitivanja faze IV i pokazuje u kojoj mjeri trenutačni status kliničkih ispitivanja faze IV ispunjava njihove zahtjeve.

1. Introduction

Adverse drug reaction (ADR) is a big health concern, responsible for more than 2 million hospitalizations each year in the U.S. alone, and ADR is associated with billions of U.S. dollars in cost every year in the developed countries. (1)
Although rigorous pre-marketing studies are required and performed for all new drugs, the safety profile of a drug at the time of regulatory approval is often incomplete due to some characteristics of phase I– III trials, such as limited sample sizes, short duration of studies, and strict inclusion/exclusion criteria.
There are usually important health issues that need to be solved at the time a product is licensed. That is shown by the fact that approximately 20% of drugs acquire new black box warnings in the post-marketing phase, and 4% of the drugs ultimately become withdrawn for safety reasons. Phase IV clinical trials have not been an integral part of the drug development process for very long yet.

The regulation for phase IV trials, as we know it today, began in the United States in the year of 1997, as the FDA required annual reporting via post-marketing commitment in the FDA Modernization Act. Since 2007, in the United States, the Food and Drug Administration (FDA) has been authorized by the Food and Drug Administration Amendment Act (FDAAA) to require post-marketing clinical trials to address safety concerns regarding a given drug.
In 2009, one billion dollars has been provided by the American Recovery and Reinvestment Act to studies that analyze and compare the effectiveness of new treatments. (2)

This shows a rebalancing of the weighting between pre-marketing and post-marketing studies, as the importance of post-marketing surveillance is perceived by both politics and society.

2. Overview of Pre-Marketing Phases

Every drug that is available on the market necessitates a well-defined and regulated protocol of steps. Those steps are defined by the countries and their responsible agencies, such as the Food and Drug Administration (FDA) in the United States and the European Medicine Agency (EMA) for the European Union. As trials can potentially pose harm to their participants, the primary priority is safety, followed by efficacy. If a drug seems useful in pre-clinical studies, consisting of examination of human material, e.g., tissue samples and determination of biomarkers, those studies become known as phase O studies, and a drug sponsor can request an investigational new drug application. After the regulatory agency approves the drug, it can be studied in clinical trials and, if no major adverse events occurred and the drug showed to be efficacious, the company can then submit a New Drug Application. (2)
The way to drug approval can be very long and exhausting, and only 20% of the drugs that started a phase I trial actually get approved on the market. (3) There are different explanations for such high failure rates.

In a review of the causes for such high failure rates, Fogel found that out of 640 phase III trials with novel therapeutics, 54% failed in clinical development. 57% of those failed to prove efficacy, which was the main cause for failure. They found that this was related not exclusively to the drug itself, but also to many intrinsic flaws in study design, sample selection, and inappropriate statistical endpoint, or simply underpowered clinical trial (i.e., sample size too small to reject the null hypothesis).(4) This study shows how crucial adequate study design is for every step of a drug's development, and although phases I-III are the most rigorously regulated in drug development, they are susceptible to flaws in study design.

After examinations in test tubes, on cell cultures, and animals, the substance will be tested on humans for the first time in a phase 1 trial; however, the study is mostly carried out on healthy people. The new active ingredient is being tested for the first time on study participants who are volunteers, under strict medical supervision, mostly in specialized facilities. The aim is to obtain preliminary safety of use and an initial presentation of the pharmacokinetic and pharmacodynamic profile of the active ingredient.

In phase II, the new drug is usually used for the first time in patients who suffer from the disease or symptoms for which the active ingredient is intended. In this phase, efficacy and safety are checked, and dose-finding studies are carried out. Typically, 100 to 500 volunteer patients take part in the studies as test subjects.

Phase III studies are clinical studies in which the drug is tested on a larger group of patients, in order to see whether effectiveness and safety can also be confirmed in many different patients.
In phase III studies, the main focus is to prove the efficacy and find out common side effects and interactions with other drugs. In order to be able to prove the significance, larger numbers of patients are required. Several thousand patients can therefore be included in phase III studies, depending on the incidence of the tested indication. (2)

3. Shortcomings of Preclinical Trials for Detecting Real World Effectiveness

Why cannot research be terminated after the preclinical trials? According to the general perception of the public, most of the important research is done at the time the drug is released on the market.
Phases I to III, due to their nature, are limited both in scope and size. Although randomized controlled trials are still the gold standard for any drug-approval process, they have many shortcomings related to their real-world effectiveness and their significance for determining a complete side effect profile of a drug, partly due to relatively small sample sizes. In phase III, which has the biggest sample size of all pre-marketing trials, the average number of participants is around 500.Also, they contain only a specific type of participants due to the exclusion of children or pregnant women, which limits generalizability. (5)
One of the key problems are the limitations that determine in how far the results from Randomized Clinical Trials (RCT) can be generalized to everyday clinical practice. Due to

the inclusion and exclusion criteria that characterize all randomized clinical studies, the sample group is much more homogenous than the rest of the general population. The results are very precise but solely on a very narrow spectrum of the population.

Inclusion criteria determine the attributes for the study population necessary to enroll in the study, such as the stage of the disease or specific pathophysiological characteristics. They usually identify a population in which it is hoped that the drug can be shown to work.

Exclusion criteria specify characteristics that exclude patients from participation and often include factors such as comorbidities or concomitant treatment, also factors that could mask the effect of the intervention. The problem arises when these criteria exclude the whole subgroup that will later receive the drug (after it gets approved), and the relevant effects of the drug in that subgroup will not be sufficiently researched; for that reason, there could be some complications.(6)

The main identified exclusion criteria in RCT are pregnant or lactating women, due protections afforded in HHS regulations. That is not because pregnant women are vulnerable to compulsion from others or are incapable of protecting their own interests, but rather because of the potential for injury to the fetus and its sovereignty.

To combat this problem, in the United States, Congress required that a task force on research specific to pregnant women and lactating women (PRGLAC) be formed according to the 21st Century Cures Act. This group will report a plan to find and address gaps in knowledge regarding safe and effective drugs for pregnant and lactating women.(6)

Another group that gets exempted from pre-marketing trials are infants, children, and adolescents due to severe tragedies that have occurred in the past, like the sulfanilamide disaster in 1937, which took the lives of more than 100 people, many of them children, and led to the development of the 1938 Federal Food, Drug, and Cosmetic Act, the statute that today remains the basis for FDA regulation of these products.(7) Another example is the thalidomide scandal of 1953, resulting in more than 10,000 children born with a range of severe deformities.

Other reasons a person can get excluded from a trial can be various comorbidities; for example, in a sample of over 30 drug trials submitted in the United states, there were more than 60 percent of exclusions based on liver aspartate aminotransferase (AST) or alanine

aminotransferase (ALT) levels, and more than 50 percent due to kidney function measured by creatinine clearance (CrCl).(6)

Also, there are other obstacles for a heterogenous study population besides exclusion criteria, like, for example, geographical distance, financial issues, or historical mistrust. The most prominent example was the notorious Tuskegee Syphilis Study that led to long-lasting tension among the African American population towards clinical research programs in the United States.

(8)

Other restrictions of RCT are the rules for concomitant medication use, which is different from polypharmacological treatment regimens in real-world settings, in which many patients take various medications at the same time for their acute and chronic conditions.

(9) Also, as a follow-up in a RCT, which is different from the normal follow-up at a local GP office, patients get monitored more closely and precisely than it would otherwise be possible. In contrast, it is characteristic of phase IV ST studies that these criteria are not as strict as in the previous studies, so the results are more applicable to the general population, a key reason why phase IV studies are so crucial. One of the biggest issues with a drug after it gets released on the market is the knowledge gap concerning its potential for rare side effects that are economically nearly impossible to determine before the drug gets released on the market, as it would require a significant increase in sample size to detect rare but severe side effects. Also, there is the possibility for lead-time bias, as certain adverse effects could have a long latency period that extends over the length of the RCT, so that complications could develop well after the release on the market.

Castle and Lewis described in their paper, “Postmarketing surveillance of adverse drug reactions,” that, to identify with 95% certainty 3 adverse drug reactions that happen at a frequency of 1/10,000, you would need a sample size of 65,000 people. This necessary sample size depends on three factors: the back-ground incidence of a reaction in the general population; “the magnitude of the drug's effect; and the required probability that a drug's adverse effect if present will be detected.”(10)

Expected incidence adverse reaction	Numbers of patients to be observed to of detect 1, 2, or 3 events		
	1	2	3
1 in 100	300	480	650
1 in 200	600	960	1300
1 in 1000	3000	4800	6500
1 in 2000	6000	9600	13000
1 in 10000	30000	48000	65000

Picture reference: Suvarna V. Phase IV of Drug Development. *Perspect Clin Res*. 2010;1(2):57-60.

As a result, the drug cannot be considered researched enough at the time it is released on the market. Also, enlarging sample sizes to detect ultra-rare side effects and to prolong phase III trials cannot be the solution, so it should be clear that post-marketing research is an essential part of every developed drug that has immense significance for the well-being of the people.

4. Post-Marketing Surveillance

4.1 Phase IV Clinical Trials

Post-marketing research, post-marketing surveillance, and pharmacovigilance were used interchangeably with phase IV studies in the most cases, as the main concern is further to examine the drug's efficacy and safety.

"However, not all FDA-mandated (classical phase IV trials) research consists of randomized controlled trials (RCTs), and not all post-marketing activities are limited to safety issues (pharmacovigilance)"(5), so that these terms need further discussion.
Post-marketing research is used to define the overall research that happens after the drug is released on the market, which can have different initiation reasons. Some are required by regulatory agencies as a part of the approval process of a drug.
This can vary from country to country. In case they are required by regulatory agencies, they are called post-marketing commitment studies (PMCs). There are also others that are started by the pharmaceutical companies or from outside investigator-initiated trials. One of the examples of investigator-initiated studies worth mentioning is the Vioxx study, which led to the removal of the anti-arthritis drug Rofecoxib. (11,12)
A common perception is that to identify rare adverse effects, phase IV studies might not be necessary, as there are worldwide spontaneous reporting systems. Most known is the MedWatch program in the United States, or EudraVigilance, the European Union data processing network and management system for reporting and evaluation of suspected adverse drug reactions.

Most information that reaches regulatory agencies about adverse drug events come from spontaneous reporting systems. They play a crucial role in post-marketing surveillance, but there are big limitations to their significance. Therefore, phase IV studies are required to assess the drug's safety profile. Adverse drug effects get noticed by a physician in the hospital or a general practitioner and then are reported to the MedWatch program, which has to deal with an increasing number of reports: from 150,000 in the year 2000 to more than 350,000 in 2009. (2) Due to the big amount of data and lack of clearly identifiable casualization, it is really difficult to rely on those reporting systems as the sole origin of

information on adverse drug events. The failure of those mechanisms could be seen in the serious cardiovascular adverse drug effects of the anti-inflammatory drugs Vioxx and Meridia: Although there was spontaneous reporting of adverse drug events to the MedWatch program, the FDA withdrew Meridia only after the European SCOUT (Sibutramine Cardiovascular Outcome Trial) study found out that patients with preexisting cardiovascular conditions, who were treated with long-term sibutramine, had an increased risk of myocardial infarction and stroke. (2)

The central problem of those spontaneous reporting systems is that they rely on voluntary reporting. This can lead to a distorted image of the actual number of adverse events, as the data might be incomplete and the spontaneous reporting is susceptible to different types of bias. For example, reporting might be increased as the public perception of a drug is more skeptical. Also, single rare adverse events might trigger fear and increased reporting of false-negative results, due to increased physician visits or physicians being more likely to report adverse events that might resemble a more severe form of adverse drug reaction. Further, those reporting systems rely on trust in the drug manufacturers' will to collect and report the data to the agencies, which could potentially harm their own financial interest.

Nevertheless, those spontaneous reporting systems have a crucial role in post-marketing surveillance, as they give the first impression of potential adverse events, even if they cannot definitely be linked to the drug, as they could in a clinical trial. Also, they serve as an important resource of inspiration for scientists to start further investigation. Phase IV trials are performed as post-marketing research further to examine the new drug regarding its safety profile, new indications, usability in different population groups, and different formulation effectiveness.

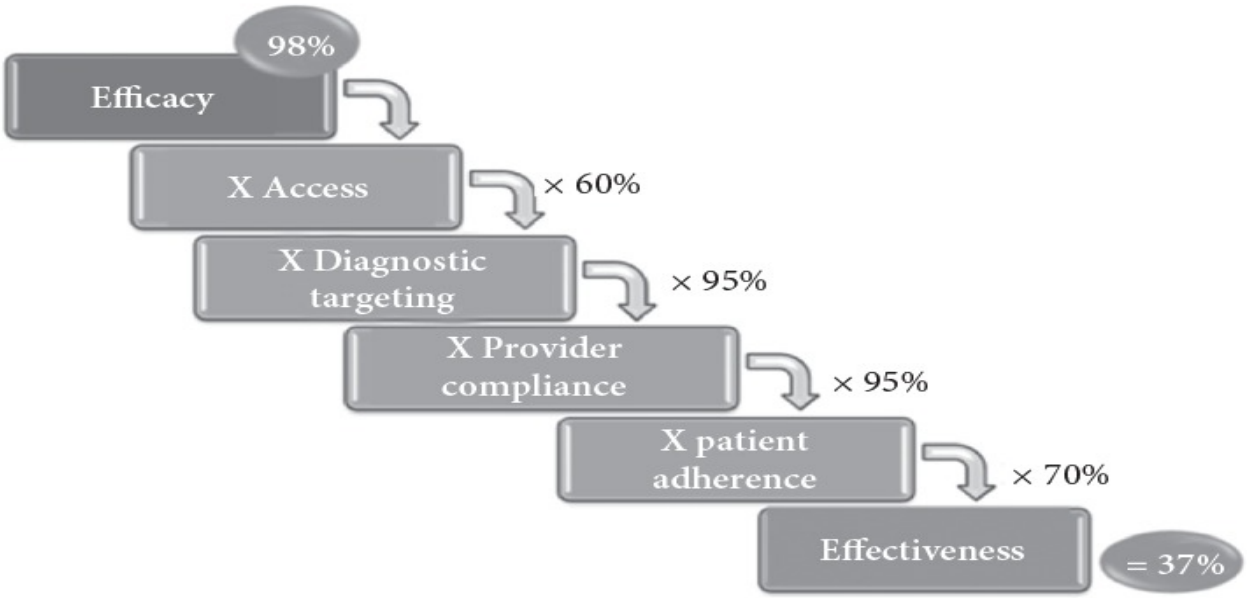
One of the main aims of phase IV studies is to detect flaws in the pharmacovigilance of a drug. Pharmacovigilance is “the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other drug-related problem” (13), and it plays a key role in ensuring that patients receive safe drugs.

Phase IV studies are mostly used in the scientific world for all post new drug application clinical development programs. Some use the term “phase IV” only for FDA-requested clinical trials, and the term “phase V” for studies that determine relevant market expansion research, interested in new indications and new formulation of the drug. (14)

There are various types of studies that can constitute a phase IV trial, some of which are observational studies, standard RCT, drug-drug interaction studies, or special population studies. They all can be requested by the FDA or the European Medicines Agency after the new drug application (NDA) approval. A key characteristic is that those are generally larger studies with a more diverse and heterogeneous population than in the preclinical studies. Also, the setting of the study is aimed to resemble real-world conditions in regular clinical care.

There are many potential goals of a phase IV trial. One of them is to understand the real-world effectiveness and external validity in contrast to efficiency and internal validity that pre-market studies are mostly concerned with. Internal validity refers to the amount of confidence that the causal relationship being examined is correct and not confounded by other factors or variables; therefore, the study design of preclinical trials and population selection has to be very selective. External validity refers to the extent to which results from a study are generalizable to other situations or groups. Therefore, the external validity is more a concern of phase IV studies, as they have the goal to examine the drug in its real setting, in contrast to the artificial setting of a phase III trial.

Similarly, phase IV trials are better defined as studies that demonstrate effectiveness, rather than efficacy. Efficacy is defined as the ability of the drug to perform under optimized conditions and to establish the desired effect. Those optimized conditions are usually established by a RCT, whereas effectiveness is characterized by the drugs ability to achieve the effect under usual clinical conditions. (15) Although this may sound like a debate over language, it has an enormous effect on the whole study design of a phase IV trial, as they choose an effectiveness approach over efficacy.



Effectiveness studies ask whether the intervention works in real-world practice setting in contrast to resource-intensive "ideal setting." They usually have a heterogeneous population with few to no exclusion criteria, and the intervention is applied with flexibility, and concurrent interventions are permitted.
Although it brings obvious benefits to assess the real-world effectiveness, it also brings problems with it, as results from effectiveness studies get less of a value attributed, since studies that pursue an efficacy approach are still regarded as more meaningful.

In their article, "Effectiveness Versus Efficacy: More Than a Debate Over Language," Fitz and Cleland showed an example of an exclusion of a clinical trial for further meta-analysis, that utilized an effectiveness approach. This could lead to less, rather than more scientific significance. It was illustrated on a set of specific
guidelines for therapeutic exercise in the management of chronic lower back pain (LBP).

The panel excluded a randomized trial by O'Sullivan. This study compared a group performing stabilizing exercises with a comparison group that received usual treatment as massage, heat, ultrasound, and exercise program. The study was not included, because the comparison group did not meet the narrow inclusion criteria of the study, whereas the other

studies comparison groups performed no treatment at all. The results of the effectiveness study by O'Sullivan were quite favorable to the group receiving the stabilization training in both short and long term.

The perspective that researchers constructing the guidelines had was that the comparison group used in the study was inappropriate because it lacked internally. From a clinical perspective, use of a usual-care comparison group may have more meaning than a no-treatment group. "Demonstrating that a new intervention is superior to a standard treatment may be more likely to change practice behavior than showing that a new intervention is superior to no treatment at all"(15)

Later performed systematic reviews, which included the study done by O'Sullivan, graded it as a high-quality study, and evaluated the stabilization exercises as useful in the therapy. Without the inclusion of this study, the panel's EBCPG recommended stretching and strengthening for patients with chronic LBP.
It appears that the panel may have missed the chance to provide better guidelines to therapists, such as the specific stabilizing program used by O'Sullivan, by excluding studies using an effectiveness approach.(15) Although this example is not a phase IV study, it shows the tension between those two approaches. Phase IV trials, as their study design resembles more the standard clinical setting than a controlled sterile environment, should get more scientific significance rather than be treated with suspicion and mistrust.

4.2 Different Phase IV Study Designs

Progressively, post-marketing trials go beyond the phase I-III safety and efficacy studies that mainly are built up by RCT. Phase IV trials ask questions about long-term safety, patient acceptance, and economics of use. The answers to those questions are not just important to the regulators, but also to other stakeholders, e.g., physicians, sponsors, and the patients

themselves. In contrast to phase III trials that usually stick to a single study design, phase IV studies require different types of designs.

My aim is to give an overview of a variety of study approaches that are relevant as phase IV study designs. There are different designs all of which have their advantages and disadvantages; it cannot be said that one design is superior to another. Therefore, it is important for researchers to choose the right methodology for their specific study. (16)
Among the important strengths of phase IV research is exposure of a broad range of participants to the drug, which will result in more "real-world" information about the drug's characteristics.

4.2.1 Post-Marketing Surveillance Studies

"Not all Phase IV studies are Post-marketing surveillance studies, but all PMS is a phase IV study" (17)
The main focus of this kind of study is to find adverse drug reactions and to guarantee proper safety monitoring of the drug after it is released on the market. It is an adjuvant to spontaneous monitoring systems and has the purpose to detect background signals, which might indicate that there is an issue.
(18)

There is a transition from the regulatory agencies taking reactive approaches towards more proactive approaches, focusing on prevention.

There was a big change in the last few years, as regulatory authorities realized that they have to apply measures to suppress the incidences of adverse drug reactions of novel drugs. Examples for this can be found in Canada and the UK.

In 2013, in the United Kingdom, around 150 of 100,000 people reported an adverse drug event. Similar were numbers in Canada, where up to 22,000 deaths as a result to novel drugs have been reported each year. (19) It is expected that the number of reported ADR accounts for only 10% of ADR, and the rest is going unnoticed by regulatory agencies. By implementation of better post-marketing surveillance standards, there should be a better overview of the actual scope of the extent of ADR. (19)

The study design of PMSS could be general, open studies with less strict regulation of exclusion criteria than pre-marketing studies, but still follow the exact indication and contraindication criteria that the drug is affirmed to. The purpose of wider criteria is to allow the study to capture effects of the drug that may have been previously hidden in phases I-III.

In certain countries, those types of studies are required for every drug that is released on the market, such as in Japan. In the United States, Periodic Safety Update Reports (PSUR) must be delivered in a predetermined interval to regulatory agencies. PSUR by law requires pharmaceutical companies to evaluate reports of side effects collected from all over the world and to check whether that results in necessity for further restrictions on use or additional side effects have to be taken into account. (20)

Nevertheless, in most countries, like in Germany, the state does not insist companies conduct post-marketing surveillance studies as a requirement for every drug.(20)

4.2.2 Large Simple Trials

This kind of trial design is a mixture of a randomized clinical trial and an observational study. (21) LST are useful to identify small or modest effect of a drug, that becomes relevant when noticed in a larger population with a certain common disease or condition. (5)

LSTs have a relatively large sample size, compared to RCT of phase I-III, and enough statistical power to detect minor treatment effects. Another advantage of LST is that, due to the large sample size, the effects of random error can be minimized. Nevertheless, LST are rarely used as phase IV study design, due to obstacles in implementing LSTs for regulatory purposes.

Regulators have made enormous progress in the implementation of large simple trials. The FDA issued a guidance in 2012 on "Determining the Extent of Safety Data Collection Needed in Late Stage Premarket and Post approval Clinical Investigations," in order to help trial sponsors determine the amount and types of safety data that should be collected.

The aim of this was to increase quality of safety assessment without undermining integrity and validity of trial results and decrease the burden on researchers and patients participating in a study, as well as to lower trial costs by facilitating the increased use of large, simple trials. (22)

A meta-analysis done by Smith et. al, which compared PROBE trials (Prospective, Randomized, Open-Label, Blinded Endpoint) design, which is a variant of LST, with double-blind trials in hypertension, showed that both study designs can be statistically equivalent. (3)

4.2.3 Randomized Clinical Trial

Similar to the study design in the previous phases, a phase IV study can be designed in a randomized clinical trial, with blinded or double-blinded study groups, a placebo group, or a comparator drug to ensure that the participants get their necessary treatment. These designs share a very similar structure and methodology like the previous phase III studies. (23)

Those trials try to evaluate further the drug's efficacy in a controlled environment and can be seen as an extension of the previous clinical trials.

However, as it is stated in Farahani's article, "Clinical data gap between phase III clinical trials (pre-marketing) and phase IV (post-marketing) studies: evaluation of etanercept in rheumatoid arthritis,"(9) there are key differences between phase III clinical trials and phase IV post-marketing studies that involve patient characteristics, the clinical setting (environment), and the manner of drug use.

They found that the rheumatoid arthritis drug profile was different between the patients receiving etanercept in the phase IV community cohort study and the patients enrolled in the RCTs. (24) As phase III studies take place in a heavily regulated environment, compared to phase IV studies, there will be always a gap between the results of both, even if they are designed similarly, as phase III is more concerned with efficacy of the drug, whereas in phase IV, the drug is tested for its real-world effectiveness.

In contrast to pre-marketing phases, phase IV clinical trials, instead of the conventional Randomized Clinical Trial (RCT), can use a more modern approach of the adaptive trial design (25), which has the goal to increase the flexibility. The main issue with widespread use of adaptive clinical trials in the setting of phase IV clinical trials is that the adaptive study design could undermine its validity and integrity.

4.2.4 Physician Experience Studies (PES)

As the name implies, those are types of studies based on physician reports and therefore often criticized for a lack of scientific accuracy. The benefit of this kind of study design is that it is relatively cheap, compared with large and complex RCT, they help physicians gain experience with the new intervention, and they provide information from a big heterogenous pool of population that helps to assess the real-world effectivity and safety profile of the drug. (3)
An example for this kind of study was the article, "Antihypertensive safety and efficacy and physician and patient satisfaction: results from a phase 4 practice-based clinical experience trial with diltiazem LA," done by Glasser. The study was performed as a large-scale, open-label study with more than 15,000 physicians and 130,000 patients enrolled. The results showed that DLA was safe and produced clinically meaningful reductions in blood pressure, as well as a high degree of physician and patient satisfaction. (26)

5. Discussion

In their study, "Overview of phase IV clinical trials for post-market drug safety surveillance,"(1) X. Zhang and Y. Zhang analyzed a data set of more than 4,000 phase IV trials between 2004 and 2014 from the ClinicalTrial.gov website. ClinicalTrial.gov is a

registry of all the trials as a prerequisite for publication. Also, due to the Food and Drug Administration Amendments Act of 2007, which set most of the guidelines for phase IV trial designs, it is required to upload trial designs and results, so they are publicly available. They analyzed only clinical trials that defined their study end point as safety only or safety and efficacy, by using logistic regression analysis. Of all the analyzed trials, just around 300 were focusing exclusively on safety, whereas the majority of phase IV trials investigated both efficacy and safety. This could be explained as safety. Although it is generally expected to be the key concern for a phase IV trial, it is more of a side investigation, as companies that sponsor most of the trials are primarily interested in further investigating the effectiveness in various fields in which the drug is used.

Common perception is that, with progression of clinical trials from I to IV, phase IV clinical trials are the largest studies, regarding the sample size. This is commonly incorrect and was documented in a study from X. Zhang, who found that the median number of participants per trial was 104 (IQR:48-258) (1) Regarding the phase IV trials that evaluated exclusively for drug safety, for which you could expect higher participant numbers, more than 60% had fewer than 300 participants enrolled, and only around 4% had large participant numbers higher than 3,000. Small participant numbers might be sufficient in evaluating the efficacy of a drug in a certain small subgroup of population to investigate safety issues with a new drug that might have gone unnoticed; a larger pool of people would have been necessary to detect adverse effects, especially those that occur less often. They also found that there were considerate differences regarding the study design, depending on who sponsored the study. Phase IV studies that were sponsored by a university or a school were more likely to be using blinding, compared to industry-sponsored studies.

Further, they found different methodologies in different specialties, as oncological trials used randomization and blinding less often, compared to cardiovascular studies. This could be a result from different expectations toward oncological drugs being more focused on efficacy, rather than safety, or because oncological drugs are used in more personalized treatment, where blinding and randomization are not appropriate.

This shows that phase IV trials are perhaps not as well-regulated as they should be, regarding the relatively small sample sizes and the heterogeneity in quality. Stricter regulation regarding size and quality could be implemented as they are for pre-marketing trials. Also, they illustrated how most studies are sponsored by industry itself (more than 50%), whereas those that are directly sponsored by U.S. federal states or the NIH (National Institute for Health)

hold the number of only less than 3%. As finding safety issues is not the primary concern of the industry, a change in sponsoring of phase IV trials might improve their quality.

The researchers also found that there is a demographic shift of phase IV studies more towards the Asian and pacific countries, as now more than 30% are placed in Asia and the Pacific. An ethical problem with clinical trials could be that demographic shifts place the study location away from the destined population, and more towards countries where the regulations are less strict.

There is a general change in the public view about the value of balance between risk and benefits of medicine, which started a debate about how to improve monitoring of drug safety after the approval of a drug. As drugs become more and more efficacious and the differences in efficacy get smaller, the focus of newly developed drugs is their adverse event profile. As most clinical trials have not kept up with this change, still the majority focuses on providing superior efficacy.

A prominent example is the Vioox trial, a phase IV trial that led to the drug's withdrawal from the market, next to that of other prominent examples, such as cerivastatin, fenfluramine, and others, due to safety concerns. That has caused a lot of criticism from various sides, e.g. members of the U.S. Congress and medical journals. (11).

In the APPROVe trial, Rofecoxib (Vioox) was compared to Naproxen (in a phase IV trial) that had the aim to discover potential new indications of Viox to cause reduced risk of recurrent adenomatous polyps. During this trial, which was not a pharmacovigilance study, the researchers found that Rofecoxib had a significant 4-fold increased risk of acute myocardial infarction. Although the Vioox phase III trial was relatively large for a clinical trial with more than 8,000 participants, it did not catch any change in cardiovascular risk, as it only showed in people taking the drug for more than 18 months, seen in the later APPROVe trial. (11)

After the Vioxx case was discussed in Congress, some FDA officials requested further withdrawal of other drugs, such as Valdecoxib or Rosuvastin. Although this was not the unitary decision of the FDA, it shows that there need to be objective criteria which regulate the withdrawal of drugs independently from public sentiment and other influences.

Some experts requested that newly developed drugs, which are released on the market, receive a probationary license, which gets renewed only after post-marketing studies have shown an adequate safety profile, where the benefits surpass the risks. In the current state of post-marketing surveillance, those changes would impose serious problems for the industry, as many phase IV trials are not finished or are closed for some reason. A big problem is the limited warranty of the regulatory agencies, as their main power is to withdraw the drug from the market, which would be out of scope as punishment for poorly performed phase IV trials. Therefore, there should be an implementation of strict requirements of phase IV trials both in scope and in timing by the regulatory agencies, which can be reinforced through financial penalties or other measures. (5)

There needs to be a balance between the growing safety concerns of the public that demands more actions taken to ensure drug safety and the economical profitability with the aim of ensuring further innovation in drug research. (11) This could be achieved by more governmental involvement in phase IV studies and more investigator-initiated trials, as they can guarantee more independence from companies' interests. For investigator-initiated trials, there are several reservations that have to be adhered to. The trial has to be requested by independent investigators, there has to be a scientific need for this study, and the company's benefit from the study has to be irrelevant. Such studies are generally regarded as more reliable, as they act relatively independently from the industry and "there is always greater weight attached to non-industry sources of data." (27)

These kinds of studies are still rare as phase IV clinical trials, since most are still initiated and performed by the industry, but they could have an enormous positive impact for both the public, due to increased safety data, and the industry as new indications would be discovered.

A common criticism is the lack of transparency of working regulatory agencies and the process of evaluating the results from Phase IV clinical trials or other post-marketing studies. There is usually a closed gremium involved, separated from outstanding scientists, deciding whether a drug should or not be removed from the market. The fact that the participants are "being bound to secrecy will inevitably lead to the impression that all these many experts agreed on the conclusion set forth by EMA", as stated by Brinth. (28) This leads to a problem for independent scientists or organizations to question the validity of the gremium's

decision, because their findings will be announced as one, and their thought processess and inner disagreements do not reach the public.

This situation is especially problematic if a topic is controversial and of high public interest. It would be very useful for the public and scientists to know whether the decision for an approval or disapproval was on the edge or voted for with a big majority. (28) Also, it could have a positive impact on the public's mistrust and anxiety for adverse drug reactions from vaccinations, if they are not presented to the public as an end result of an international organization. Compared to this, in the United States, there are academic/non-profit organizations like RADAR (Research on Adverse Drug Events and Reports) and Public Citizen that play a role in pharmacovigilance and pose a counterpart to the powerful Food and Drug Administration. For example, the activist group Public Citizen criticized an "inappropriate close collaboration" between the FDA and Biogen regarding the pharma's controversial Alzheimer's disease drug Aducanumab and the related FDA advisory committee documents. (29) The review from a group of outside experts panned the drug, saying that it didn't pass muster for an approval. But the official FDA documents, released several days before the review, said something different. This shows a more equal balance of power and the controlling functions those non-governmental organizations have on the drug release. A similar non-governmental authority in the European Union that can question the EMA results would be desirable. Similar European organizations with a comparable scope and significance are mostly missing.

6. Conclusion

Phase IV clinical trials are of enormous importance for the public, as spontaneous reporting systems often face underreporting, and serious adverse drug events can go unnoticed in pre-marketing trials. The various study designs of phase IV studies, compared to phase I-III trials, should not be a limitation to their significance, but rather be their strength as they examine the

drug in a real-world setting. There is a need for stricter regulation, predetermined requirements, and control of phase IV studies, as there is great variance from study to study.

Currently, there is an optional and lightly regulated study system of phase IV studies, compared to much more rigorously controlled phase I-III studies. Also, changes in laws that regulate and guarantee increased quality and standard of phase IV trials and regulatory agencies have enough power to force those commitments.

Higher regulation means higher expenses for drug development as the standards increase. This could negatively affect the initiative of drug development, as the monetary expenses for new drugs are enormous. Higher standards and more regulation of phase IV trials would potentially worsen the problem.

There could be a rebalancing of expenses and effort, as it is proposed by some experts, by shifting the focus from pre-marketing trials more towards post-marketing trials. That is shown by the fact that 10% of news drugs, released on the market in the United States, either were withdrawn or received an additional warning of serious or life-threatening side effects after their release on the market. (5)

Many phase IV trials have to increase sample sizes to be able to find rarer ADE. Thus, a shift from phase IV trial designs that are similarly designed like pre-marketing studies in a RCT design focusing on internal validity, towards larger observational large simple trials would be desirable. They may lack some internal validity but are able to show better information about the drugs' real effectiveness in the open market and can help with assessing a proper safety profile of the drug.

7. Bibliography

1. Zhang X, Zhang Y, Ye X, Guo X, Zhang T, He J. Overview of phase IV clinical trials for postmarket drug safety surveillance: a status report from the ClinicalTrials.gov registry. Available from: http://dx.doi.org/10.1136/bmjopen-2015-010643
2. Umscheid CA, Margolis DJ, Grossman CE. Key concepts of clinical trials: A narrative review. Postgrad Med. 2011;123(5):194–204.
3. Glasser SP, Salas M, Delzell E. Importance and challenges of studying marketed drugs: What is a phase IV study? common clinical research designs, registries, and self-reporting systems [Internet]. Vol. 47, Journal of Clinical Pharmacology. J Clin

Pharmacol; 2007 [cited 2021 Apr 2]. p. 1074–86. Available from: https://pubmed.ncbi.nlm.nih.gov/17766697/
4. Fogel DB. Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: A review. Vol. 11, Contemporary Clinical Trials Communications. 2018. p. 156–64.
5. Glasser SP, Salas M, Delzell E. Importance and challenges of studying marketed drugs: What is a phase IV study? common clinical research designs, registries, and self-reporting systems. J Clin Pharmacol. 2007;47(9):1074–86.
6. US Food and Drug Administration. Public Workshop: Evaluating Inclusion and Exclusion Criteria in Clinical Trials IN CLINICAL TRIALS. 2018;(July). Available from: https://www.fda.gov/media/134754/download
7. Ballentine C. Sulfanilamide Disaster. FDA Consum Mag [Internet]. 1981;(June 1937). Available from: http://www.fda.gov/AboutFDA/WhatWeDo/History/ProductRegulation/SulfanilamideDisaster/ucm2007257.htm
8. Charlotte Paul, MB, PhD, and Barbara Brookes P. The Rationalization of Unethical Research: RevisionistAccounts of the Tuskegee Syphilis Study and theNew Zealand "Unfortunate Experiment."
9. Kremer JM, Weinblatt ME, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Jackson CG, et al. Etanercept Added to Background Methotrexate Therapy in Patients With Rheumatoid Arthritis Continued Observations. ARTHRITIS Rheum. 2003;48(6):1493–9.
10. Castle WM, Lewis JA. Postmarketing surveillance of adverse drug reactions. [Internet]. Vol. 288, British medical journal (Clinical research ed.). Br Med J (Clin Res Ed); 1984 [cited 2021 Apr 9]. p. 1458–9. Available from: https://pubmed.ncbi.nlm.nih.gov/6426604/
11. Station W, Food US, Vioxx T, Congress US, Oath TH, Chan A, et al. Drug safety on trial. Nature. 2005;434(7033):545.
12. Importance and challenges of studying marketed drugs: what is a phase IV study? Common clinical research designs, registries, and self-reporting systems - PubMed [Internet]. [cited 2021 Apr 2]. Available from: https://pubmed.ncbi.nlm.nih.gov/17766697/
13. Härmark L, Van Grootheest AC. Pharmacovigilance: Methods, recent developments and future perspectives [Internet]. Vol. 64, European Journal of Clinical Pharmacology. Eur J Clin Pharmacol; 2008 [cited 2021 Feb 25]. p. 743–52. Available from: https://pubmed.ncbi.nlm.nih.gov/18523760/
14. Grudzinskas C. Design of Clinical Development Programs. In: Principles of Clinical Pharmacology. Elsevier Inc.; 2007. p. 501–17.
15. Fritz JM, Cleland J. Effectiveness Versus Efficacy: More Than a Debate Over Language [Internet]. 2003. Available from: www.jospt.org
16. Jaeger BT, Deshpande R. Choosing a Phase IV Study Design. Pharm Mark. 2003;29–30.
17. Suvarna V. Phase IV of Drug Development. Perspect Clin Res [Internet]. 2010 Apr [cited 2021 Apr 9];1(2):57–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21829783
18. Cvmp TTO, et al. Committee for Veterinary Medicinal Products Guideline for the Conduct of Post-Marketing Surveillance Studies of Veterinary Medicinal Products. Practice. 2000;(April).
19. Raj N, Fernandes S, Charyulu NR, Dubey A, G. S. R, Hebbar S. Postmarket surveillance: a review on key aspects and measures on the effective functioning in the context of

the United Kingdom and Canada. Ther Adv Drug Saf [Internet]. 2019 Jan [cited 2021 Apr 17];10:204209861986541. Available from: /pmc/articles/PMC6661791/
20. Periodic safety update reports (PSURs) | European Medicines Agency [Internet]. [cited 2021 Feb 25]. Available from: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/periodic-safety-update-reports-psurs
21. Suvarna V. Phase IV of Drug Development. Perspect Clin Res [Internet]. 2010 Apr [cited 2021 Feb 24];1(2):57–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21829783
22. The Imperative of Overcoming Barriers to the Conduct of Large, Simple Trials [Internet]. 2014 [cited 2021 Apr 18]. Available from: http://www.fda.gov/downloads/Drugs
23. Studies PI V, Management L. 24. 2018;25:383–91.
24. Kremer JM, Weinblatt ME, Bankhurst AD, Bulpitt KJ, Fleischmann RM, Jackson CG, et al. Etanercept added to background methotrexate therapy in patients with rheumatoid arthritis: Continued observations. Vol. 48, Arthritis and Rheumatism. 2003. p. 1493–9.
25. Mahajan R, Gupta K. Adaptive design clinical trials: Methodology, challenges and prospect [Internet]. Vol. 42, Indian Journal of Pharmacology. Wolters Kluwer -- Medknow Publications; 2010 [cited 2021 May 9]. p. 201–7. Available from: /pmc/articles/PMC2941608/
26. Glasser SP. Antihypertensive safety and efficacy and physicians and patient satisfaction: Results from a phase 4 practice-based clinical experience trial with diltiazem LA. Adv Ther [Internet]. 2006 Mar [cited 2021 Apr 18];23(2):284–94. Available from: https://pubmed.ncbi.nlm.nih.gov/16751161/
27. Suvarna V. Investigator initiated trials (IITs). Perspect Clin Res. 2012;3(4):119.
28. Brinth BL. Responsum to Assessment Report on HPV-vaccines released by EMA November 26 th 2015. 2015;
29. Public Citizen calls out “inappropriate close collaboration” between FDA and Biogen, urges investigation | FierceBiotech [Internet]. [cited 2021 Feb 25]. Available from: https://www.fiercebiotech.com/biotech/public-citizen-calls-out-inappropriate-close-collaboration-between-fda-and-biogen-urges

8. Biography

I was born in Saarbrücken, a town in the southwest of Germany, on October 22nd, 1996. There, I spent most of my childhood and completed high school at the Ludwigsgymnasium. Besides school, I spend my free time doing mainly sports activities, like basketball and water polo, but also developed a passion for chess in my high school years. After my graduation, I

decided to return to my Balkan roots and started studying medicine in Zagreb.