University of Twente.

Using multi-criteria decision analysis to evaluate the potential of biosimilars to lower costs in oncology

Hilde Tuenter, Bsc.

h.tuenter@student.utwente.nl

DEPARTMENT OF HEALTH TECHNOLOGY & SERVICES RESEARCH

EXAMINATION COMMITTEE

Dr. J.M. Hummel, University of Twente
Professor Dr. M.J. IJzerman, University of Twente
Professor Dr. D.W. Raisch, University of New Mexico

Preface

I proudly present to you my dissertation. Over the past six months, I have been working on this dissertation. This whole process was informative, interesting and above all challenging. At the beginning, I had no knowledge regarding biosimilars and their potential. After these past six months, I can conclude that I have gained a lot of knowledge about this subject. In addition, I became aware of the fact that the global economic impact of cancer is tremendous, and the need for alternatives to reduce this economic impact is high. The aim of this dissertation was to contribute to the scientific knowledge with regard to biosimilars in oncology. The future will tell, however, if biosimilars are able to fulfill their potential to lower costs in oncology globally.

I could not have done this survey all by myself. For that reason, I want to thank some people. First of all, I want to thank professor dr. Raisch, professor dr. IJzerman and dr. Hummel. They gave me the opportunity to study abroad. This experience enriched me as a student as well as a human being. Besides, their insight and feedback have helped me to improve the quality of my dissertation. Furthermore, I want to thank Henk Broekhuizen and Shikhar Shrestha for their assistance with the data analysis. Lastly, I want to thank my family and friends for supporting me during this process and their help whenever I needed it.

Abstract

Background

Even though biosimilars have the potential to bend down the cost curve of oncology by at least 10 percent, their actual uptake appeared to be less evident than expected. Currently, both patients and healthcare providers are not confident about the efficacy, safety and interchangeability of biosimilars compared with their reference product. The limited knowledge regarding biosimilars in oncology may have an impact on the attitudes and perceptions of stakeholders in oncology. To evaluate the potential of biosimilars to lower costs in oncology, it is important to determine the preferences of stakeholders in oncology for biosimilars.

Objective

A multi-criteria decision analysis is designed to determine which factors stakeholders consider to be important with regard to the uptake of biosimilars in the Dutch oncology setting, to prioritize the role of post-marketing studies.

Methods

An online questionnaire is used to reveal the preferences of Dutch stakeholders about biosimilars in cancer care. The included stakeholder groups consisted of physicians, oncologists, pharmacists, employees of health insurance companies involved in formulary decisions or benefit structures, (government) policy makers and researchers. A combination of ciscrete choice experiments (DCE) and the analytic hierarchy process (AHP) is used to determine the biosimilar preferences. Pairwise comparisons are established to obtain relative weights and overall priorities of four criteria, i.e. costs savings and three factors that relate to equivalence with the reference biological: effectiveness, safety and immunogenicity. An additive model is used to see if providing additional information to individuals has an effect on preferring biosimilars. A logistic regression model was fitted to investigate if this potential effect is important and significant. Subgroup analyses are performed to investigate if preferences differ per subgroup.

Results

A total of 34 respondents completed the questionnaire, of whom 7 had a baseline preference that very strongly or extremely favored the biosimilar. Adding information contributes to an increase of the preference, baseline score. This increase in preference score applies for all decision criteria. Providing all post-marketing data along with approval data, and all post-marketing data along with approval data as well as additional cost savings, is associated with significant higher odds ($p<0.01$) of preferring biosimilars over biologics. It is also observed that post-marketing effectiveness data along with approval data was associated with significantly higher odds ($p<0.05$) of favoring biosimilars over biologics.

A total of 23 respondents met the consistency ratio threshold of >0.20 and were included for the analysis of the pairwise comparisons. When the four decision criteria are compared with each other, post-marketing safety data is considered to have the highest relative importance to the respondents with regard to the uptake of biosimilars (weight = 0.37). After post-marketing safety data, post-marketing effectiveness data was considered to be the most important (weight = 0.323), followed by post-marketing immunogenicity data (weight = 0.204) and cost savings (weight = 0.104).

Conclusions

With reference to the multi-criteria decision analysis the factors are elicited that are most important with regard to the uptake of biosimilars in the Dutch oncology setting. Stakeholders in oncology prefer post-marketing effectiveness data and post-marketing safety data along with approval data in the case of biosimilars in oncology. The combination of all post-marketing data along with approval data will most likely result in preferring biosimilars over its reference biological. Post-marketing studies will play a major role in the potential uptake of biosimilars in oncology, and are required before their implementation on a large scale can be realized.

Table of Contents

List of figures

List of tables

Abbreviations and definitions

ACS	American Cancer Society
ACSCAN	American Cancer Society Cancer Action Network
AHP	Analytic Hierarchy Process
BIO	Biotechnology Industry Organization
CKD	Chronic Kidney Disease
CML	Chronic Myeloid Leukemia
DCE	Discrete choice experiment
EMA	European Medicines Association
ESA	Erythropoiesis Stimulating Agent
EPO	Epoetin alfa
FDA	Food and Drug Administration
G-CSF	Granulocyte colony-stimulating factor
HVEGF	Human vascular endothelial growth factor
IMP	Investigational medicinal product
Incidence	"The total number of people that develop a particular disease/faced a particular health-related event within a specified period of time"[103]
MCDA	Multi-Criteria Decision Analysis
MDS	Myelodysplastic Syndromes
Mortality (rate)	"The total number of people that died within a specified period of time, or due to a particular cause"[104]
NCI	National Cancer Institute
PBSC	Peripheral blood stem cell
PML	Progressive Multifocal Leukoencephalopathy
PRCA	Pure Red Cell Aplasia
QoL	Quality of Life
SMD	Small molecule drug
WHO	World Health Organization

1. Introduction

1.1 Background

1.1.1 Impact of cancer

Cancer is one of the leading causes of morbidity and mortality in the world. According to the World Health Organization (WHO), the incidence of cancer was approximately 14 million in 2012 worldwide. The global incidence of developing cancer differs between males and females. For males, there are 205 new cancer patients per 100,000 males worldwide. The incidence is slightly different for females, because there are globally 165 new cancer patients per 100,000 females[1]. This can be represented by a male:female incidence ratio of 10 versus 9, respectively[2]. In addition, approximately 8.2 million people died as a result of cancer in the same year. Of every 100,000 males, 126 men died due to cancer. In comparison, there were 83 deaths for every 100,000 females[1]. These results imply that there is a male:female cancer mortality ratio of 10 versus 8[2]. The expectation, however, is that the incidence will increase with approximately 70 percent in the years ahead, resulting in a worldwide incidence of almost 24 million cancer patients a year in the next twenty years[3]. This increase in incidence is shown graphically in figure 1[4]. Assuming that the mortality rates remain the same, this expected increase will consequently result in a global increase of cancer related deaths.

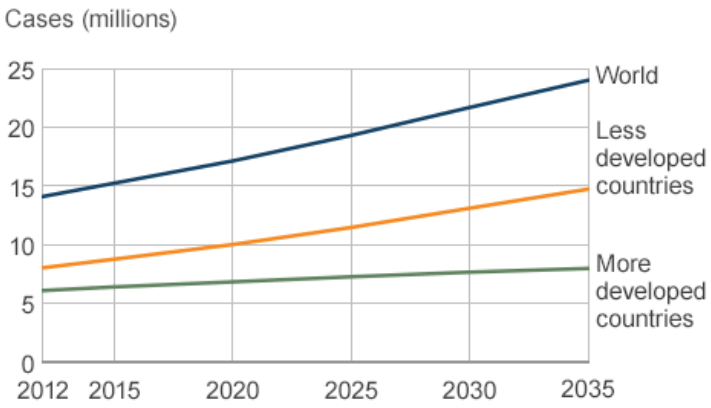


Figure 1 *Predicted global cancer cases*

Besides the fact that cancer is one of the leading causes of death among human beings, cancer has also the "most devastating economic impact in the world"[5]. In 2008, the global economic impact of cancer was $895 billion. This impact is a result of the indirect mortality costs, which can be calculated by the loss of productivity due to the disabilities and the premature deaths of cancer patients. Heart diseases have the second most devastating global impact, with a total loss of $753 billion[5]. From this, it can be concluded that the economic impact of cancer is almost 20 percent higher. It is important to note that this economic impact does not even include the direct costs for cancer care. Including this costs would logically increase the total economic impact of cancer even more. To exemplify this: in 2010 the total costs of direct cancer care were approximately $125 billion in the United States alone[6]. The expectation is, however, that the United States has to deal with a rise of the costs of direct cancer care to an amount of $173 billion in 2020. This means that the costs of direct cancer care will increase with almost 40 percent within one decade[7].

Several reasons can be designated for this substantial increase of the costs of direct cancer care. First of all, the United States has, like many other countries, to deal with both an ageing and growing population. Therefore, the prediction is that the prevalence of cancer will increase. Another important reason can be found in the fact that new technologies and treatments are developed, which are more expensive compared with the current cancer therapies[6]. The last decades, small molecule drugs (or 'chemical' drugs) were used in particular to prevent and to treat many diseases worldwide. Since the development of new technologies and treatments, small molecule drugs (SMDs) were partly replaced by biological products that had entered the market. A major and expanding role was and still is reserved for these so-called 'biologics'. However, the development didn't end after the introduction of these biologics. In recent years it was found that it is possible to develop drugs that are said to be equivalent to those biologics in terms of efficacy, safety and immunogenicity, at lower costs. These drugs are also known as 'biosimilars'.

In this dissertation, it will be investigated if these biosimilars indeed might have the potential to reduce costs, particularly in the field of oncology. Before it is possible to provide an answer to this statement, it is essential to compare them with products that currently are used to treat cancer. Therefore, both small molecule drugs, biologic drugs and biosimilars will be discussed.

1.1.2 Small molecule and biologic drugs

Before the introduction of biologics, small molecule drugs were seen as the medicines to prevent and treat many diseases in the world. Their existence had a large impact on improving public health as well as increasing the life expectancy of human beings. For example, in 1900 almost 33 percent of the fatalities in the United States were related to tuberculosis, diarrhea and pneumonia. Currently, these three causes of mortality are rare due to the fact that it is possible to prevent or to treat them with small molecule drugs. In addition, the existence of small molecule drugs resulted, among others, in an increase of the life expectancy of almost 30 years within one century (50 years in 1900 up to almost 77 years in 2000)[10]. It can be stated that the introduction and usage of small molecule drugs had a significant impact on both the life expectancy and the public health as a whole.

Although the substantial impact of small biologic drugs, biologics became the most important therapies to treat complex, life-threatening diseases in the last decade[9]. Their development fundamentally changed the treatments of, for example, diabetes, rheumatoid arthritis and anemia[11-13]. The development of biologics also positively contributed to treatment of cancer. A biological therapy can encourage the immune system of the human body to attack the existing cancer cells. To accomplish this, biologics in the form of both vaccines and bacteria can be used. Thus, this usage of biologics, which is popularly referred to as immunotherapy, acts against the existing cancer cells in an indirect way through the immune system[8]. The use of biologics can stimulate the way the immune system is responding to the existing cancer treatment, resulting in "stopping, controlling or diminishing the process that allows the growth of cancer cells"[9]. Biologics can also offer the possibility to treat cancer in a direct way. In that case, the biologics will disturb the molecules which are affecting the growth of the tumor itself.

The use of (monoclonal) antibodies, cytokines and recombinant DNA products are examples of such 'targeted therapies'[8]. The fact that biological therapies can treat cancer directly by primarily focusing on the cancer cells in the human body, makes them particularly accurate and personalized. Consequently, less healthy cells will be affected by the use of these biological therapies. The possibility that a patient experiences side effects as a result of the cancer treatment they received, can therefore be reduced by

using biologics. Thus, the goal of biologics can be to destroy the cancer cells, or to minimize potential side effects of cancer therapies such as chemotherapy[9].

When small molecule, chemical, drugs are compared with biologics, several differences can be described. The most important difference between these medicines can be found in the way they are manufactured. Small molecule drugs consist of chemicals and are developed through chemical synthesis. That implies that drugs are made "by combining specific chemical ingredients in an ordered process"[14,15]. Biologics are, however, manufactured using components derived from living organisms[17]. Their main differences are summarized in table 1:

Small molecule drugs	Biologic drugs
Made by chemical synthesis	Made by living cells
Defined structure	Heterogeneous structure Mixtures of related molecules
Easy to characterize	Difficult to characterize
Relatively stable	Variable Sensitive to environmental conditions
Usually taken orally	Usually injected
Often prescribed by a general practitioner	Usually prescribed by specialist Immunogenicity

Table 1 *Differences between chemical and biologic drugs[16]*

1.2 Problem definition

1.2.1 Biologics and their concerns

1.2.1.1 Costs

One of the reasons of the increase of health care costs is the rise of the cost of cancer therapies. In the United States, about 70 percent of the sales of anticancer drugs relate to products which have been developed and introduced only in the previous 10 years. Many of these anticancer drugs are biologics or biological treatments. Biologicals are extremely expensive because of, among others, their entitlement to patent protection and their complex manufacturing process. Within a year from now, biological treatments will cover five out of the ten most expensive medication expenses. With regard to cancer treatment, 40 percent of the therapies consists of biologic drug treatments. This 40 percent is equivalent to a total of $100 billion in drug sales worldwide[18]. Besides, the manufacturing of biologic drugs is far more expensive than the manufacturing of SMDs. The reason for this difference can be found in the fact that, because of its complexity, the fixed production- and facility costs, and the costs of the required clinical trials are much higher[19].

The costs of cancer care are, partly due to biologics, an important reason for the ever-increasing costs of health care in a majority of the countries in the world. In the United States alone, the increasing treatment costs per individual cancer patient are the main reason for the increased health care costs of the whole country[20]. Not uncommon are cost-effectiveness ratios that exceed the thresholds which are widely accepted ("$20.000 up to $30.000 per QALY in the UK, US$50.000 up to US$100.000 per QALY in the US")[21]. Besides that, it is predicted that the spending on health care will increase with 6.2% annually up to the year 2018 in the United States, due to both the ageing population and the predicted increasing prevalence of cancer[6]. This implies that in 2018 the total amount of money spend on health care will be approximately $4.4 trillion in the US[22]. Since it is predicted that the prevalence of cancer will increase, it

can be expected that the labor force will decline since a greater proportion of the population will suffer from cancer. Consequently, there will be a loss of productivity and an increase in (in)direct mortality costs. The expected increase in incidence of cancer, in combination with the rising costs of both the loss of productivity, the costs of (direct) cancer treatments and the total amount of money spend on health care, will result in a non-sustainable trend[23].

1.2.1.2 Safety and immunogenicity concerns

Biologics might be more accurate and personalized than small molecule drugs, their complexity makes it more difficult to develop them. Even a small modification during the manufacturing process could seriously affect the structure of the drug, resulting in both potential safety and effectiveness issues for every individual patient[15]. Modifications in the structure of the drug could trigger the immune system to attack it, because the substance might be recognized as being foreign to the human body. This response is also referred to as 'immunogenicity'. Immunogenicity often results in tachyphylaxis, which signifies that the efficacy of the product decreases[24]. The reason behind this is that the antibodies the human body creates during the immune response, could ensure that the drug won't be effective in any further intake[9]. The possibility exists that patients develop an allergic reaction towards the natural proteins that their own body produced, which could even worsen their situation and health state[25]. The goal of a biologic is to stimulate cells to produce proteins, which in turn should attack the cancer cells. Differences within the structure of these proteins could, however, induce unwanted or unforeseen cell behavior, altered effectiveness and insolubility[26]. As a consequence, these cells could produce different molecules or proteins which are not interchangeable to the proteins that should have originated. When the drugs are not equivalent to the original biologic product, different clinical outcomes may arise between patients that take the drug[9]. Currently, however, no procedure exists that shows the effectiveness and safety of the biologic drugs in advance[15].

In the last decades, several problems with biologics were reported regarding safety and immunogenicity issues. Examples which show the consequences that modifications can have during the manufacturing process, are Epogen and Eprex. These biologics were prescribed to patients that suffered from anemia, which means that there is a low number of red blood cells in the patients' blood. Anemia can be a result of kidney failure, or it can appear in cancer patient as an adverse reaction of, for instance, chemotherapy[27]. Both Epogen and Eprex were made from erythropoietin and the same technology was used, but their manufacturing process slightly differed. This difference expressed itself through very diverse clinical outcomes in the patients that received them[28]. It was noted that both drugs could lead to pure red cell aplasia (PRCA), which means that the patients become allergic to the protein epoetin that the human body produces itself. The prescription of Epogen resulted in 5 of these cases between 1998 and 2004. The use of Eprex, however, resulted in 175 cases of this severe adverse reaction in the same time period[28]. This substantial difference between both biologics might have been caused by the difference in the manufacturing process, which clarifies that even the smallest modifications can have a tremendous impact on clinical outcomes.

There are more examples to mention which indicate that safety concerns regarding biologics are not without reason. The biologic Bentuximab Vedotin (BV) appeared to be effective as a treatment for, among others, relapsed Hodgkin lymphoma[29]. According to a research of Gandhi and this colleagues, however, there is a chance that patients receiving this biological treatment could develop pancreatitis[30]. The use of Bentuximab could also potentially result in progressive multifocal leukoencephalopathy

(PML), which is a 'virus-induced central nervous system infection'[31]. It was found that five patients, who suffered from lymphoid malignancies and were treated with Bentuximab, developed PML after a certain period of time. Four of these five patients died because of this adverse reaction[31]. Although this research only consisted of five patients, the seriousness of the adverse reaction should increase the awareness among clinicians, physicians and the pharmaceutical industry.

Even though only two examples of adverse reactions are described, it is beyond question that the manufacturing process of biologics is of great importance for the clinical effects it can have on patients. The manufacturing process is, however, the most difficult part due to the use of living cells. When the use of biologics in (cancer) care continues, health care professionals should take into account their potential to result in safety and immunogenicity implications. Biologics are, however, in their turn more accurate and personalized when compared with small molecule drugs, and they changed the treatments of life-threatening diseases fundamentally. If biologics should be used in order to deliver the best health care possible, the patients that receive them should be monitored closely to reduce the severity of potential adverse events.

1.3 Potential solution

1.3.1 Biosimilars

1.3.1.1 Bending the cost curve

The use of biologics has several benefits, but their usage is incredibly expensive and differences in the manufacturing process may cause serious side effects. The fact that a number of biologics will reach their patent expiration date at a relatively short notice, can offer the opportunity for other drug therapies. A new method to manufacture drugs that might be promising, could be the introduction and implementation of copied versions of these biologics. These copies are also referred to as biosimilars, or "similar biological medicinal products"[32]. The Food and Drug Administration, abbreviated the FDA, defines biosimilarity as follows: "the biological product is highly similar to the reference product notwithstanding minor differences in clinically inactive components"[33]. In addition, biosimilarity implies that "there are no clinically meaningful differences between the biological product and the reference product in terms of safety, purity, and potency of the product"[33]. Thus, a biosimilar can be seen as a biologic product that is approved and of which the quality, safety and efficacy can be compared to the reference product[34].

The main reason why it is said that biosimilars are promising, can be found in the fact that it has the potential to reduce the costs of oncology as well as the health care as a whole. Several studies concluded that the implementation of biosimilars could result in total costs savings of the health care in the United States by $1 up to even $108 billion. Andrew W. Mulcahy and his colleagues performed a literature review regarding the potential cost savings of biosimilars and summarized all these findings into a table[35], which is presented on the next page. According to this literature review, the use of biosimilars could result in reduced unit prices of 10 up to even 50 percent.

Study	Approach	Scope	Time Frame	Price reduction	Savings
Grabowski et al., 2007 as applied in Goodman et al., 2009 (base case)	Economic model	6 major categories of biologics, top 20 biologics by sales only, all payers	2009-2019	12% to 20%, varies by product	$10 billion (2.4% of baseline spending)
Grabowski et al., 2007 as applied in Goodman et al., 2009 (sensitivity analyses)	Economic model	6 major categories of biologics, top 20 biologics by sales only, all payers	2009-2019	12% to 40%, varies by product	$1 billion to $44 billion (0.2% to 10.5% of baseline spending)
Ahlstrom et al., 2007 (Avalere Health)	Actuarial model	Federal payers only	2008-2017	10% to 51%, varies by product and increasing over time	$3.6 billion (0.6% of baseline spending)
Engel and Novitt, 2007	Actuarial model	Excludes Enhanced Primary Care, Medicare Part B only (office-based, physician-administered biologics)	2007-2016	Unknown	$14.4 billion
Miller and Houts, 2007 (Express Scripts)	Actuarial model	Select markets, all commercial payers	2007-2016	25%	$71 billion (baseline not reported)
CBO, 2008	Actuarial model	All biologics	2009-2018	20% to 40%, varies per product and increasing over time	$25 billion (baseline not reported), $7 billion of which accrues to the federal government
Shapiro et al., 2008	Actuarial model	Top 12 biologic classes	2010-2019	25% to 35%, varies by assumption	$67 billion to $108 billion

Table 2 *Estimated savings of biosimilars in the United States*[35]

It is expected that by 2018 many biologicals will reach their patent expiration date in both Europe and the United States. The total amount of money these biologicals are worth in sales lies between $64 and $67 billion worldwide. A few of these biologicals are developed to treat patients who suffer from cancer. One study predicts that in Europe approximately €1.6 billion can be saved annually in direct cancer treatment costs, when biosimilars will replace a number of the biological drugs that will expire by 2016[20]. Besides, the costs of biosimilars can be 10-51 percent lower in comparison with their reference biological drug[36]. Within 20 years, the amount of money saved in pharmacy drugs in the United States can reach up to $378 billion[20]. Therefore the assumption is made that biosimilars can be seen as a potential solution to bend down the cost curve of oncology and health care as a whole.

These potential estimated savings on national level are interesting for the government, health insurance- and investing companies. On patient level, the implementation of biosimilars can also be very valuable with regard to the financial access to healthcare. The fact that biosimilars are less expensive in

comparison with biologicals enables cancer patients, who could not afford the biological drug, to receive a comparable, less expensive treatment option[9]. As a result, more patients have access to the biosimilar version of the drug. It is stated that biosimilars are similar with regard to quality, safety and efficacy when they are compared to their reference product. Consequently, when an increased number of patients have access to the biosimilar due to lower costs, an improvement in public health and health outcomes could be realized. Eventually, positive effects in terms of cost-effectiveness could be established. In addition, the fact that biosimilars are less expensive results in a reduction of opportunity costs[17]. This reduction can, for example, provide patients with the possibility to obtain additional cancer treatments to enlarge the chance of survival, or result in a better quality of life.

1.3.1.2 *Approval of biosimilars*

In 2007, biosimilars were used for the first time in Europe and India[18]. Since then, an increase in the number of countries where biosimilars entered the market can be noticed. Besides several countries in Europe, many countries in Asia such as Japan, China and South Korea adopted biosimilars in their pharmaceutical industry. One very important player on the market was missing: the United States. Between 2007 and 2014, the European Medicines Association (EMA) approved a total of 21 biosimilars. A number of these biosimilars were primarily developed to treat cancer. The FDA approved its first biosimilar, Zarxio ®, in the United States by 2014[18]. Therefrom, it can be seen that there is a difference in the approval of biosimilars among the organizations which are responsible for this process. An example that demonstrates this difference is the time a biologic drug, the so-called reference product, needs to be approved before the EMA and the FDA will allow biosimilars on the market. According to the EMA, a sufficient time period to allow biosimilars on the market is when a biologic is approved for at least 10 years. The FDA, however, considers a time period of at least 12 years after approval to be appropriate. Second, where the FDA did not yet decide on allowing interchangeability –"the automatic substitution of an innovator product for a follow-on product"[37] - the EMA decided not to evaluate it at all[38]. Besides, the introduction of the European Public Assessment Report requires pharmaceutical companies to publicly report the findings regarding the comparability of the biosimilar with its reference product. There is no such regulatory requirement in the United States. As a result, chances are that patients and their physicians are not aware if any manufacturing modifications have been carried out[39]. In summary, it should be noted that the policy regarding the approval of biosimilars differs between different instances and different countries.

In the United States, the use of biosimilars is blocked by the presence of obstacles which are of regulatory and legal manner. An obstacle of great importance is the fact that clinical trials are required to include larger sample sizes compared with generics[18]. If one looks at the biosimilars that are approved by the EMA, it is apparent that the evidence that was used for their approval can be considered as limited. In general, most of the biosimilars are approved after a relatively small number of studies had been performed, ranging from a minimum of only two to a maximum of eight studies. The biosimilar Follitropin alfa, for instance, is approved after two studies were performed concerning only the primary indication of the drug[40]. In addition, these studies did not include large sample sizes. The sample sizes ranged from a minimum of 51 patients up to a maximum of 922 patients. With reference to frequent and common problems regarding the safety and immunogenicity of a drug, these sample sizes are powerful enough. However, to identify potential uncommon problems that can be associated with the use of a particular drug, a sample size of 300 up to 3000 patients is necessary[41]. With a mean sample size of 351 participants among the studies included, it is improbable that uncommon safety and immunogenicity

problems would and will be detected. A requirement for approving biosimilars is that there is at least one study performed, which compares the biosimilars with the biologic with regard to equivalence. These studies must be able to demonstrate the equivalence in terms of safety, efficacy, immunogenicity, pharmacokinetics and –dynamics[18,23]. When these requirements for approval of biosimilars are taken into account, it should be emphasized that the EMA does comply with these required conditions before they approved the biosimilars. It can, however, be stated that a smaller number of patients were treated before approval of biosimilars, compared with the number of patients that were treated before biologic therapies were approved.

In addition, there is indistinctness regarding the regulations of biosimilars. In 2012, the FDA did compose a so-called draft guidance with regard to the uptake of biosimilars. There were and are, however, still many disagreements about the adoption of regulatory standards among the members of the FDA[37]. There is also a notable variation in the uptake of biosimilars among the countries that have approved their usage, partly due to the way their health system is organized. The way countries reimburse and incentivize the use of biosimilars, in combination with the existence of variation between the different health care institutions, resulted in diverse outcomes between and within countries[38]. Besides, health care professionals are not aware when to prescribe a biosimilar drug, because of the fact that the policy regarding the prescription of biosimilars is still very limited. The remaining uncertainty and the presence of these differences results in serious delays in both the development of the market and the approval of biosimilars[42].

1.3.1.3 Existing concerns regarding biosimilars

Even though biosimilars have the potential to bend down the cost curve of oncology by at least 10 percent[36], their actual uptake appeared to be less evident than expected[43-45]. This can be explained by the fact that both patients and healthcare providers are not confident about the efficacy, safety and interchangeability of biosimilars compared with their reference product[46]. Similar to biologics, biosimilars are developed using components which are derived from living organisms. This indicates that their manufacturing process is highly complex. The difficulty of producing biosimilars makes it a relatively expensive undertaking for pharmaceutical companies. Even though biosimilars have the potential to result in cost savings, pharmaceutical companies may decide not to take the risk if it is uncertain whether or not they will be used. Furthermore, biosimilars are no exact duplications of biologics due to modifications and differences in the (environment of the) host cell. Already conducted research with regard to biologics has shown that even a small modification during the manufacturing process can cause severe adverse events in (cancer) patients. Variability in the manufacturing process is, however, inevitable due to its complexity. Thus, biosimilars will never be perfectly equivalent to their reference products. It should, however, be noted that the same applies for the production of biologics. No single production of a biologic is perfectly the same to the previous produced biologic. This does not necessarily imply that its quality is inferior or superior to the previous one[42], but in some cases (manufacturing) differences did result in serious adverse events.

The production of biosimilars is inextricably linked to complexity and variability. The fact that there might be differences in the development and, consequently, potential variations in the structure of biosimilars results in the presence of concerns among physicians, researchers and decision makers. According to M. Weise and her colleagues, oncologists in particular are reticent in prescribing biosimilars to their patients[42]. Although the FDA states that a drug can be called a biosimilar only when it does not differ meaningfully from its reference drugs, and the process regarding the approval of biosimilars is strict

and critical[47], stakeholders are not confident that they will have similar characteristics[48]. Currently, the number of studies that provide post-marketing data regarding effectiveness, safety and immunogenicity of biosimilars in oncology is limited. Although pharmaceutical companies face more data requirements in comparison with generic drugs, stakeholders remain concerned. Solely randomized controlled trials (RCTs) are able to determine the actual efficacy and safety of biosimilars, but post-marketing RCTs are currently rarely performed. The EMA and the FDA, however, underscore the necessity for post-marketing studies and post-marketing surveillance regarding biosimilarity[24]. Before stakeholders will be convinced that biosimilars are comparable with or perform even better than biologics, more scientific research should be performed. In addition, it should be noted that the preferences with regard to the uptake of biosimilars might differ between the different stakeholder groups, due to their various interests. To exemplify this, it can be expected that health insurance companies prefer that biosimilars are implemented on a short-term, because of their potential for cost savings. Physicians, however, are likely to have more concerns regarding their efficacy and safety. The preferences of pharmaceutical companies about biosimilars will depend on their financial interest whether or not to implement biosimilars in the market. For that reason, it is interesting to investigate if the preferences regarding biosimilars do differ among different stakeholder groups.

In summary, there are reasons why biosimilars are not used as intensively as possible. Currently, post-marketing data regarding biosimilars is limited. In addition, there are notable differences between the approval of the EMA and FDA and the policy for prescribing biosimilars is also restricted. Besides, stakeholders, and in particular oncologists, have concerns about the equivalence of biosimilars compared with the originator, biologic drug. Biosimilars, however, do have interesting advantages. The main advantages are that it has the potential to lower the ever-increasing costs in oncology and that it could increase the possibilities for patients to receive affordable (cancer) treatment. The actual role that biosimilars might play in oncology and the health care system as a whole is, however, utterly dependent on how clinicians assess them. Thus, before a significant uptake of biosimilars can be expected, a contribution to the current scientific knowledge regarding stakeholder preferences is required.

1.4 Research question

Although biosimilars have the potential to lower costs in oncology and could enlarge the possibility for patients to receive treatment options that are less expensive, several questions remain unanswered regarding whether or not stakeholders are willing to use biosimilars and how they will evaluate their (potential) effects. If the patient is interested in the use of biosimilars, he or she will still be dependent on the choices of their physicians, pharmacists and payers such as health care insurance companies[9]. If these stakeholders remain uncertain about effectiveness, safety and immunogenicity of biosimilars, or just prefer the biologic drug in relation to brand loyalty, a bright future for biosimilars cannot be expected.

Prior to this research a pilot study has been performed. The outcomes of this pilot study were, mainly because of its small sample size, not generalizable. This implies, in combination with the uncertainty about biosimilars in oncology according to stakeholders, that (additional) research is needed to investigate what evidence and (post-marketing) data are important before stakeholders are possibly willing to switch from original biologics to biosimilar drugs. Since both the similarity and the concerns of stakeholders relate to the effectiveness, safety and immunogenicity of biosimilars, the aim of this research is to investigate the importance of these factors, including cost savings, according to stakeholders in oncology.

Thus, by means of this research it is aimed to elaborate on the distinction between both the intentions as well as the revealed and elicited preferences of stakeholders. To investigate this distinction, this research will focus upon providing an answer on the following research question:

"Which preferences do stakeholders in oncology have about cost savings and post-marketing data in terms of effectiveness, safety and immunogenicity with regard to the uptake of biosimilars in the Netherlands?"

1.4.1 Sub questions

In order to provide an answer to the research question, several sub questions are formulated:

1. What is known in the current scientific literature about the equivalence of biosimilars in oncology compared to their reference biological in terms of effectiveness, safety and immunogenicity?
2. Which preferences do stakeholders in oncology have regarding the use of biosimilars compared with the use of original biologics in terms of effectiveness, safety, immunogenicity and costs?
3. What are the differences in preferences regarding the use of biosimilars in oncology in terms of effectiveness, safety, immunogenicity and costs when different biosimilars are taken into account?
4. What are the differences in preferences regarding the use of biosimilars in oncology in terms of effectiveness, safety, immunogenicity and costs when different stakeholder groups are taken into account?
5. What are the differences in preferences regarding the use of biosimilars in terms of effectiveness, safety, immunogenicity and costs when compared with the pilot study?

2. Multi-Criteria Decision Analysis

In order to provide an answer to the research question of this dissertation, a Multi-Criteria Decision Analysis (MCDA) will be performed. In this chapter the rationale of the MCDA and its value in health care decision making will be described.

2.1 Background of MCDA

Before a new drug will be marketed, there are several actions that need to be preceded. In fact, there is a variety of decisions that policy makers and regulatory authorities need to make before the new drug will be available in clinical practice. It is difficult to make these decisions, because in many cases they are very complex, versatile and they might even conflict with each other[49,50]. During this process of consideration, decision-makers have to take a large number of different criteria into account. These criteria include, for instance, the cost-effectiveness of the new drug, evidence on both the benefits and the risks of using the new drug, the disease severity and the context in which the drug will be used[51]. Besides, marketing and using a new drug involves a variety of different stakeholder groups, among which perspectives, opinions and interest are likely to differ. It is said that decision-makers tend to focus upon single criteria, where they actually have to deal with a large number of criteria and stakeholder interests at the same time[52]. To be able to make the best decision with reference to all the available, complex information, an approach is required which "integrates all factors considered by decision makers in practice, spanning clinical, economic, social, organizational, ethical, and legal dimensions"[53].

Multi-Criteria Decision Analysis, or MCDA, is an approach that is able to support decision-makers to convert the variety of complex, conflicting criteria into a comprehensive, simplified representation of these criteria. MCDA is namely an analytical method that offers the possibility to help decision-makers evaluating a number of alternatives, while taking into account different performance criteria[54]. A great advantage of MCDA is, therefore, that it makes the decision process more transparent. Pharmaceutical companies that are developing new drugs need both qualitative and quantitative information about the relative value the new product might have to decision-makers. This information can help them to design an accurate development plan and marketing strategy[55]. MCDA offers the possibility to obtain quantitative information about the potential value of new drugs. On the basis of this information a benefit-risk assessment of the new drug can be performed. This information can be of great value for the decision making of pharmaceutical companies whether or not to (continue to) develop the new drug[56,57]. A decision-making process which includes the MCDA method, consists of various aspects. The first aspect relates to the alternatives among which the decision needs to be taken. Second, the performance criteria are needed, among which the alternatives need to be evaluated. Third, the respondents needs to value the performance criteria of the different alternatives, to obtain a score that reflects these values. According to these scores, one is able to calculate the weighted score per criteria. This enables the potential to compare the different criteria and alternatives with each other with regard to their perceived importance and value[58]. To execute the MCDA method properly, the criteria and information provided in the method need to be "accessible, differentiable, abstractable, understandable, verifiable, measurable, refinable and usable"[59].

Currently, the MCDA approach that is used the most, is the weighted sum approach[58]. The idea behind this approach is comparable with the general usage of the MDCA. First, different scales will be constructed, which represent the preferences for the number of alternatives. The next step is to weight these scales, to obtain their so-called relative importance. This is necessary, because of the fact that the

relative importance of the various criteria is likely to differ between different decision makers and different countries[55]. With reference to the relative importance, the weighted averages can be calculated. These weighted averages represent the final weight for every single alternative[60]. Thus, the weighted averages of each alternative show the degree of preference for that particular alternative. A strong preference for an alternative will be reflected by a higher score, whereas alternatives with a small preference will logically have a lower score[52].

2.2 Pairwise comparison

MCDA is a term that encompasses a plurality of different weighting elicitation techniques. As stated before, using weighting within MCDA offers the possibility to determine the priorities of respondents with regard to the different performance, decision criteria. This implies that it is very important to choose the weighting technique that is able to distinguish the most important criteria from those that are assumed not to be of great importance. The ability of a technique to determine the priority of the different criteria is also referred to as the 'discriminative power'[61]. The weighting techniques that are most commonly used are (1) the five point rating exercise, (2) the best worst scaling, (3) the pairwise comparison and (4) the ranking exercise. According to the study of van Til and her colleagues (2014), all four techniques have the ability to discriminate the criteria according to their perceived importance. The pairwise comparison, however, had the highest discriminative power of the four options. Therefore, pairwise comparison is said to be of greatest value in prioritizing different criteria and alternatives, when higher discrimination of criteria is required. In addition, almost three quarters of the respondents (74 percent) indicated that they preferred using the pairwise comparison method[61]. Besides, using pairwise comparisons is said to be one of the better ways to identify respondents' preferences[62]. Although this are the conclusions of only two studies, pairwise comparison can be seen as a valuable technique for the elicitation of weights.

According to Saaty and his colleagues (2011), something can be called a judgement if two different components are compared with each other. The (cardinal) pairwise comparison utilizes these judgements, because respondents are asked to compare two different performance criteria on a numerical scale, which is reciprocal. Thus, using this scale enables the possibility to convert judgements into numerical values[63], making decision making a mathematical knowledge[64]. The ratio scale offers the respondent 17 different possibilities to answer the question, ranging from extremely preferring criteria A (a score of 9) to extremely preferring criteria B (also a score of 9). When both criteria are perceived to be of equal importance, the respondent has the possibility to answer with a score of 1. Figure 2 clarifies this 1-9 reciprocal ratio scale, by providing a schematic illustration[65]:

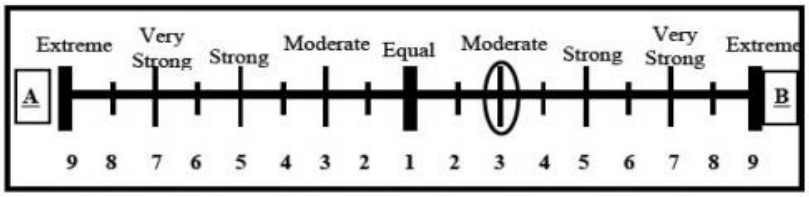


Figure 2 *Pairwise Comparison numerical reciprocal scale[65]*

The meaning of each score is provided in the following table:

Value	Meaning
1	Equally preferred
3	Moderately preferred
5	Strongly preferred
7	Very strongly preferred
9	Extremely preferred
2,4,6,8	Intermediate values

Table 3 *Ratio scale in pairwise comparison*[66]

To investigate the theoretical and empirical validation of this scale, various studies are performed. One can see articles published by Saaty[67,68], among others, for examples that ratify and validate the usage of this 1-9 reciprocal scale.

2.3 Analytic Hierarchy Process

The Analytic Hierarchy Process (AHP) is one of the most commonly used MCDA techniques, when multiple criteria are involved in the process of decision-making. This technique is already applied in a various number of studies and proved to be of great value in helping decision-makers during complex conditions[69]. The reason for its wide use, is that this quantitative technique has several advantages. One of the main advantages of the AHP is that it nearly always (in approximately 75 percent of the cases) uses the pairwise comparison technique[60]. Other advantages of the AHP are that (1) it is easy to use, which results in a user-friendly technique, (2) it is valuable when the preferences of respondents are likely to vary widely, (3) it can help decision-makers when clinical evidence is not (yet) available and (4) it uses a consistency test to correct for any inconsistent answers[54,69,70]. However, one is not only full of praise about the AHP in complex decision making. A critique is, for instance, that there is a risk of so-called 'rank reversal'. Rank reversal can occur when performance or decision criteria are added or removed, potentially resulting in making different decisions[71]. In addition, pairwise comparisons are said to be arbitrary and highly subjective. These latter points, however, have been convincingly refuted[60]. Despite the presence of this criticism, researchers and decision-makers remain continuously interested in the use of the AHP. Therefore, it can be argued that the advantages of this technique outweigh its critiques.

There are several steps that need to be followed before decisions can be made with regard to the AHP technique. These steps are conceived by Thomas L. Saaty (2008), who is also the inventor of this theory. The first step is to describe both the problem and the knowledge that is required to be able to make the decision. This step is followed by composing the decision hierarchy. The summit of this hierarchy shows the purpose of the decision. The layers below represent the decision criteria and the decision alternatives. With regard to this hierarchy, the pairwise comparisons can be drafted. Respondents are asked to compare the decision criteria in pairs. With reference to these comparisons, one is able to determine the relative importance of one criteria as compared to the other. According to the obtained relative importance of all the decision criteria, a pairwise comparison matrix can be established[72]. Figure 3 illustrates how such a matrix looks like[73].

With respect to the goal	X	Y	Z
X	1	9	3
Y	1/9	1	1/3
Z	1/3	3	1

Figure 3 *Example of a pairwise comparison matrix*[62]

From this matrix, it can be seen that three different decision criteria are compared with each other. With respect to the goal, criteria X is assumed to be 9 times (or extremely) more important than criteria Y, and 3 times (or moderately) more important than criteria Z. Besides, criteria Z is assumed to be 3 times more important than criteria Y. However, these numerical values don't provide any information about the relative weights of the three criteria compared to each other. To calculate the relative weights of each of the items, the (normalized) eigenvector, or priority vector, can be used. This vector offers the possibility to numerically rank the alternatives with respect to their perceived preferences. The vector can also "reflect intensity or cardinal preference as indicated by the ratios of the numerical values"[74]. To be able to compose the most preferred decision with reference to the numerical values and weights as a result of the group decision making process, it is required to calculate the geometric mean. It is shown that the geometric mean is a mathematical way to translate multiple individual judgements into one judgment, which represents the overall set of preferences. The geometric mean makes it also possible to convert all the different individual ordinal preferences into a comprehensive ordinal group preference[72].

Before reliable conclusions can be drawn, it is necessary to investigate if the obtained judgments are consistent. To clarify consistency in the setting of AHP, an example will be provided. Suppose that criteria A is perceived to be 3 times more important than criteria B, and that criteria B is found to be 3 times more important than criteria C. A consistent pairwise comparison between criteria A and criteria C should be that criteria A is perceived to be 6 times more important than criteria C. However, it is not uncommon that respondents make inconsistent judgments[62]. Therefore, the AHP includes a so-called consistency test which is able to correct for the presence of inconsistent answers. If the level of inconsistency is perceived to be high, it is important to revise or remove those judgments. It is said that the consistency ratio must be lower than 0.10, before it is acceptable to include the judgments into the decision-making process[69,70].

2.4 Application of MCDA and AHP in health care

Multi-criteria decision analysis and the analytic hierarchy process are already applied within the decision-making process in various application fields. The technique is, for instance, frequently used in the public administration setting. The Department of Defense of the United States makes, for instance, extensive use of the AHP in allocating their resources among various operations. In addition, the AHP is also used by British Airways in the 90s of the previous century for selecting the provider of the entertainment system in their airplanes. The AHP is even applied by motor companies like Ford to increase the contentment of their clients, and in the sporting business in the decision whether or not to retain (baseball) players[72]. In contrast, ten years ago the usage of this decision-making technique was still limited in the health care setting[52]. The usage of the analytic hierarchy process has, however, increased during the years. A literature review performed in 2011 reviewed a total number of 93 articles that focused upon the usage of the AHP technique in the health care setting[54]. Therefrom it can be stated that the application of MCDA in the health care setting increased substantially in a relatively short period of time. Its usage is found to be

valuable during complex decision-making processes, which might increase the uptake of MCDA in the context of health care even more.

In order to demonstrate the value of MCDA and AHP in health care, a study will be described in which these techniques were used to support the process of decision-making. This study, performed by Hummel and her colleagues (2013), investigated the preferences among Dutch men and women with regard to screening techniques for colorectal cancer[75]. With reference to the results, it was aimed to increase the intention to attend at the screening program among Dutch citizens. One can see the published article by Hummel and her colleagues[75] for the findings. The results of this study showed the preferences of Dutch citizens with regard to colorectal cancer screening. When one takes these preferences into account, the intention to attend to this screening program could potentially increase. For that reason, the findings of this study can be valuable in the decision-making of the technique to choose in this colorectal cancer screening program. This is only one example of a study that indicates that using multi-criteria decision analysis has the potential to numerically show preferences of relevant stakeholders to enlarge the transparency. Using both MCDA and AHP can therefore be valuable in the decision-making process in the variable, complex health care setting.

2.5 MCDA and biosimilars

It is questionable whether or not biosimilars will be marketed widely in the (near) future. In fact, the case of biosimilars can be seen as a complex health care issue in which multi-criteria need to be considered. Although chances are that developing biosimilars is profitable for the pharmaceutical industry, e.g., their development costs remain high and the usage of biosimilars is not guaranteed. Therefore, it is risky for these companies to enter the market. In addition, there might be interesting advantages of the implementation of biosimilars in health care, there are also concerns that are not insignificant. Both these benefits and risks can be regarded as a number of decision criteria. With reference to these criteria, it is possible to establish pairwise comparisons. Thus, it is conceivable to perform a multi-criteria decision analysis with the analytic hierarchy process in order to determine the potency of biosimilars. On the basis of these techniques, one is able to determine preferences among various stakeholders involved with the adoption of these newly-manufactured drugs. By means of this decision-making technique it is possible to inform stakeholders about the uptake of biosimilars while it is still in its early stages.

Currently, MCDA has not yet been applied to evaluate the potential of biosimilars. At this moment there is only a very small number of studies that even focused upon tracking down opinions of stakeholders regarding biosimilars. One of the few studies that actually aimed to determine opinions of stakeholders about biosimilars, only included rheumatologists to fill in the questionnaire[76]. Thus, it can be stated that the current available literature concerning stakeholders' opinions about biosimilars in oncology is limited. Since the debate on the adoption of biosimilars continues, a multi-criteria decision analysis including various stakeholder groups might be valuable to actually evaluate the potential of biosimilars on a large scale. Performing a MCDA is not only useful to elicit the preferences of stakeholders, it also provides valuable (quantitative) information regarding the various decision criteria. Thus, using MCDA enables one to enlarge the transparency. Besides, the results of a MCDA can be valuable for the pharmaceutical industry whether or not to develop biosimilars and which marketing strategy to use. Finally, it can positively contribute to the current scientific knowledge.

3. The pilot study

3.1 Background

Prior to this research, a pilot study is performed in which a MCDA is used to determine preferences of biosimilars regarding their uptake. In this study, use was made of a case example, followed by preference questions, swing weighting questions and demographic questions. Questions regarding the use of biologic drugs and potential experiences with adverse events were also asked, to determine if these experiences could possibly have had an impact on the answers given. With reference to these results, one was able to establish the pairwise comparison matrix and to calculate the geometric mean, or the 'mean weights'. The respondents were identified by the members of the study team. This process continued until a convenience sample size was identified. The final target population consisted of respondents from the Netherlands (N=24), the United States (N=21) and Germany (N=24). The respondents were asked to fill in an online questionnaire, which was prepared in LimeSurvey. The link to this questionnaire was provided by e-mail. The preference questions made use of pairwise comparisons. Respondents were asked to fill in the extent to which they preferred to use a reference biological or a biosimilar. For each of these pairwise comparisons they were asked to take into account other decision criteria (costs, effectiveness, safety and immunogenicity). By means of the swing weighting questions it was aimed to determine what modifications in the decision criteria would result in switching from the reference biological to the biosimilar, followed by assigning weights to these modified criteria. On the basis of the obtained results and the answers given on the demographic questions, a subgroup analysis was performed.

3.2 Results

From the 69 respondents identified, a total of 21 respondents (accounting for 43 percent) actually filled in the online survey. This study population consisted of 11 respondents from the United States and 10 respondents from the Netherlands. From the obtained results it could be concluded that a number of respondents (N=6) had provided illogical or inconsistent answers. For that reason, these results were excluded from the analyses. As a result, the answers of a total number of 15 respondents were included in the final analyses. According to these results and the calculated geometric means, the pairwise comparison matrix was developed. The findings of this matrix are shown in table 4 below:

	Cost	Effectiveness	Safety	Immunogenicity
Cost	1.00	0.84	0.88	0.85
Effectiveness	1.20	1.00	1.02	0.98
Safety	1.13	0.98	1.00	1.00
Immunogenicity	1.17	1.02	1.00	1.00

Table 4 *Pairwise comparison matrix of the pilot study*

The highest geometric mean was found in the pairwise comparison between costs and effectiveness. According to the findings, the respondents considered a post-marketing study on effectiveness to be 1.20 times more important than increasing cost savings from 25 up to 50 percent. In fact, this matrix shows that post-marketing studies on effectiveness, safety and immunogenicity are all assumed to be more important than additional cost savings. These findings are in accordance with the findings of the swing weighting questions. With reference to these results, respondents indicated that there are more likely to switch from a reference biological to a biosimilar when post-marketing data is available about biosimilars regarding effectiveness (E), safety (S) and immunogenicity (I). However, when compared to the pairwise comparisons, the results of the swing weighting show that post-marketing studies about safety and

immunogenicity are preferred over a post-marketing study regarding effectiveness. Additional cost savings (C) was assumed to be the least important criteria to switch from reference biologics to biosimilars. These results can be found in figure 4.

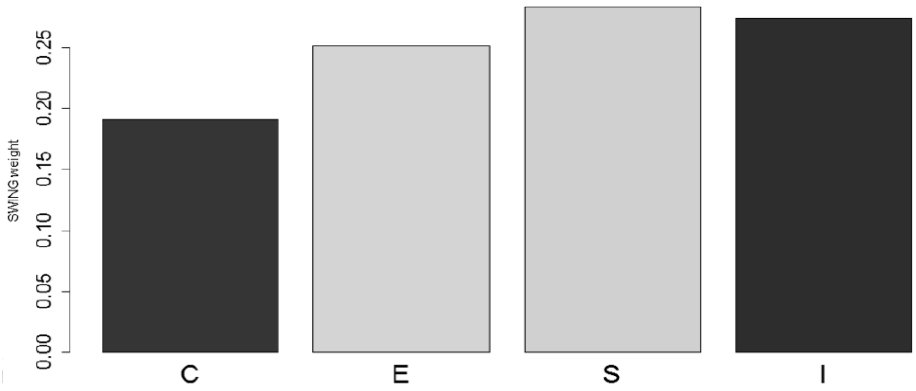


Figure 4 *Swing weights of criteria*

A subgroup analysis was performed for both the preference questions and the swing weighting questions. The results of the analysis of the swing weighting questions will be presented and explained briefly. The findings can be found in table 7 below:

Group (n)	Cost	Effectiveness	Safety	Immunogenicity
Practitioner (n=10)	0.20	0.23	0.27	0.30
Payer/policy maker (n=1)	0.23	0.31	0.31	0.15
Researcher (n=7)	0.20	0.24	0.26	0.30
Pharma (n=3)	0.15	0.29	0.35	0.21

Table *5 Swing weights subgroup analysis*

What stands out, and what was to be expected, is that the preferences differ among the subgroups. The higher the score, the greater a preference for a post-marketing study for a certain decision criteria. It is noteworthy that the additional cost savings were not assessed as most important by any subgroup. Even the subgroup that represents the health payers and policy makers preferred post-marketing studies of effectiveness and safety over additional cost savings. It should, however, be noted that this subgroup consisted of only one respondent.

The role in health care was not the only aspect among which the respondents differed. Since respondents from both the United States and the Netherlands completed the survey, also a subgroup analysis was performed regarding the country of origin. According to this analysis, the Dutch respondents were found to have a significantly greater preference ($p=0.03$) for the biosimilars in comparison with the respondents from the United States (respectively 12.09 ± 3.10 versus 8.45 ± 4.53). An explanation might be that the EMA approved a larger number of biosimilars, whereby the possibility exists that the Dutch respondents had more experience with biosimilars. Another explanation could be the related to the differences between the organization and incentives of both the health care systems.

3.3 Limitations

Upon completion of the pilot study it can be concluded that there are some limitations. The first limitation relates to the fact that only 21 respondents completed the survey, of which only 15 appeared to be useful for further analyses. Thus, the sample size of the pilot study was small. Another limitation relates to the means of identification. The members of the study team used the convenience sampling technique and identified potential respondents of whom it was assumed that the likelihood of participation was high. The risk of selection bias is, based on this fact, considerable. Besides, this sampling technique could result in both under- or over-representation of different subgroups. In fact, this actually occurred, since one subgroup only existed of one respondent. Therefore, it can be stated that it is difficult to obtain answers which are generalizable.

In addition, there were some limitations regarding the case example and the questions posed. The case example provided a patient case, in which the respondent needed to consider treating a patient with filgrastim. The biosimilar of filgrastim is approved in Europe. In contrast, in the United States biosimilar filgrastim was not approved at the time the pilot study was conducted. This could imply that the respondents from Europe might have had positive or negative experiences with this biosimilar, whereas respondents from the United States did not had any previous experience with biosimilar filgrastim. Differences between both continents could therefore, at least partly, be explained by the choice of this drug. Furthermore, it was found that respondents experienced problems related to the swing weighting questions. The underlying reason could be that it was unclear to the respondent what was expected. Thus, chances are that the obtained answers do not represent the initial preferences of the respondents. Besides, not all the decision criteria are plotted against each other in the pairwise comparison questions. For that reason, it was not possible to calculate some of the geometric means. This resulted in the fact that the pairwise comparison matrix could not be drawn completely, with the consequence that results are inconclusive.

With reference to the findings and limitations of pilot study, modifications are applied to improve the quality of the current survey. These modifications mainly concern the case example, the means of identification of the respondents and the questions asked in the survey. For example, the case example will provide a drug of which the biosimilar is not approved yet to be sure the respondent does not have any experience with the biosimilar of the reference biological. In addition, since the respondents experienced difficulties with answering the swing weighting questions, these questions will be removed from the current questionnaire. Furthermore, the AHP will be applied to draft the pairwise comparisons to be able to draft a complete pairwise comparison matrix.

4. Methods

4.1 Study population

To provide an answer on the research question, Dutch stakeholders in oncology are approached to complete an online questionnaire. The included stakeholder groups, the target population, consisted of physicians, oncologists, pharmacists, employees of health insurance companies involved in formulary decisions or benefit structures, (government) policy makers and researchers. Both men and women are included, aged between <30 and >60 years in May 2015. The initial email with the request to participate to this survey was sent to Dutch hospitals, health insurance companies, pharmaceutical companies, and instances and companies that are concerned with cancer research and policy making. The snowball sampling method was used. Using this method enabled the possibility that identified members of the target population referred the survey among other people with the same characteristics that are in interest of the study[77]. By using the snowball sampling method it was aimed to obtain a sample size that is representative for the target population.

It was found that previous experience with weighting techniques could result into less distinction or discrimination between the criteria[61]. For that reason, people who already participated to the pilot study were excluded. If judgments of respondents appeared to be repetitive or inconsistent (>0.20), judgments were revised or excluded from the analysis.

4.2 Study design

A systematic literature review is conducted with regard to biosimilars, as disclosed in the relevant scientific literature. The objective of this literature review was to determine the current scientific knowledge of biosimilars in oncology. To investigate the preferences and opinions regarding the potential role of biosimilars in oncology, a combination of the Analytic Hierarchy Process (AHP), a Multi Criteria Decision Analysis technique, and Discrete Choice Experiments (DCE) was used to evaluate and estimate the preferences of the target population. Pairwise comparisons are established with reference to the findings of the literature review. The pairwise comparisons allowed respondents to quantitatively judge the different decision criteria. LimeSurvey is used to create an electronic, web-based questionnaire to identify the preferences. The questionnaire consisted of preference questions, pairwise comparisons and demographic questions and can be found in Appendix A. The target population received an email containing a link to access the questionnaire. Two weeks after the initial email, a reminder was sent to non-responders. After the commencement, the survey was open for two months. By means of the questionnaire, data is obtained to answer sub questions 2-5 and the research question.

4.3 Objectives

4.3.1 Decision criteria

To determine and evaluate the preferences of the study population about biosimilars in oncology, several outcomes needed to be obtained with the questionnaire. The decision criteria, or key variables, that are measured relate to the factors that stakeholders consider to be the most important when evaluating the potential of biosimilars, compared to the original biologics. The decision criteria used were identified with reference to the literature search on biosimilar pharmaceuticals in oncology:

- Efficacy: The ability of the biologic or biosimilar to achieve the desired effect.
- Safety: The usage of a biologic or biosimilar must imply that the patient is free from the occurrence of risk of injury, danger, or loss.

- Immunogenicity: The characteristic of the biologic or biosimilar to enable an immune response, or the extent to which the biologic/biosimilar has this possibility.
- Cost savings: The amount of money saved as a result of using biosimilars instead of biologics.

4.3.2 Decision alternatives

In this survey, respondents had to take two decision alternatives into account. The first decision alternative regards biologics, the drugs that are currently most often used in the treatment of life-threatening diseases like cancer. The other decision alternative concerns biosimilars:

- Biologics: Can stimulate the way the immune system is responding to the existing cancer treatment, resulting in "stopping, controlling or diminishing the process that allows the growth of cancer cells"[9,14].
- Biosimilars: A biologic product that is approved and of which the quality, safety and efficacy can be compared to the reference, biological product.

Figure 5 below shows the decision hierarchy, with the goal, the decision criteria and the decision alternatives:

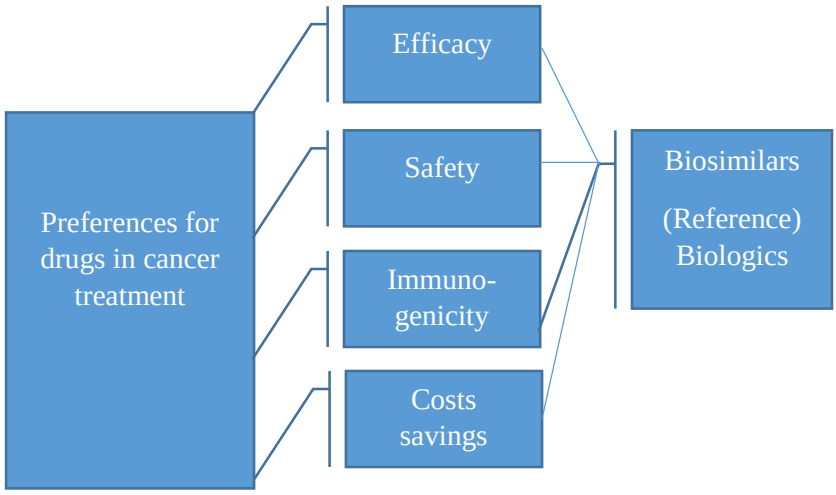


Figure 5 *Analytic hierarchy process structure including decision criteria and alternatives*

4.3.3 Other objectives

Because of the fact that different stakeholder groups are included within the study, the baseline, primary preferences among the subgroups are compared. It is also investigated if the initial preferences changed, when the cost savings were more substantial. Besides, it is aimed to investigate if the preferences in terms of effectiveness, safety, immunogenicity and costs vary among various stakeholder groups, and if these preferences change if they have had positive or negative experiences with biologic products. For that reason, the following objectives are necessary to be able to make the comparisons:

- Demographic information (age, gender, profession)
- Previous experience (with either biologics or biosimilars)

4.4 Acquisition of preferences

To identify the preferences of the target population, an online questionnaire was used. This questionnaire started with an introduction, where some general facts regarding biologics and biosimilars were presented. Thereon, an example of an AHP pairwise comparison was provided. This example demonstrated how to score the 1-9 reciprocal, ordinal scale in the way that corresponds with the preference of the respondent. After this example, a general question was asked in order to determine the baseline preference of biosimilars. Before the actual preference questions started, the respondents had to choose between five reference biological products. The product chosen must represent the biological which the respondent most often prescribed for cancer treatment, or the biological that the respondent was most familiar with. Since an attempt was made to approach different stakeholders in oncology, chances were that if one single reference biological product was chosen, some of the respondents were not familiar with that particular product. Therefore, respondents could choose between the five most prescribed biological drugs according to sales in both Europe and the United States[92]. In addition, it was important to provide reference biological drugs of which no biosimilar was available at the time of the survey. This could, namely, result in different judgments as a result of previous experiences. As a result, the five biological products were (1) CD20 antigen inhibitor, e.g. Rituximab, (2) human epidermal growth factor receptor 2 inhibitor, e.g. Trastuzumib, (3) human vascular endothelial growth factor inhibitor, e.g. Bevacizumab, (4) tyrosine kinase inhibitor, e.g. Imantinib, and (5) epidermal growth factor receptor inhibitor, e.g. Cetuximab. The product selected was used in the description of the case example. After the case example, the respondents were asked to fill in the actual questions that must represent their preferences towards both the decision criteria and the decision alternatives (figure 2). While answering these questions, the respondent needed to take the chosen biological product into account.

In the questionnaire the assumption is made that post-marketing studies show that the biosimilar is equivalent to its reference biological product. Furthermore, it is important to note that the questionnaire consists of three different sections. In the first section, six questions were asked to determine the role of cost savings and post-marketing studies on the extent to which either the reference biological or the biosimilar is preferred. These questions aim at finding a threshold value for which people are likely to switch to a biosimilar product. Instead of making a comparison between a pair of decision criteria, the respondent needed to indicate, on the 1-9 reciprocal scale, to what extent either the reference biological or the biosimilar is preferred when different decision criteria are added on top of approval data. Figure 6 below provides an schematic representation of the way these questions are established. As shown in figure 6, these questions relate to both single and combinations of decision criteria alternatives. One could say that the way these questions are drafted relate more to a combination of AHP questions and discrete choice experiment questions. For that reason, the analysis of these questions differs from the actual AHP analysis.

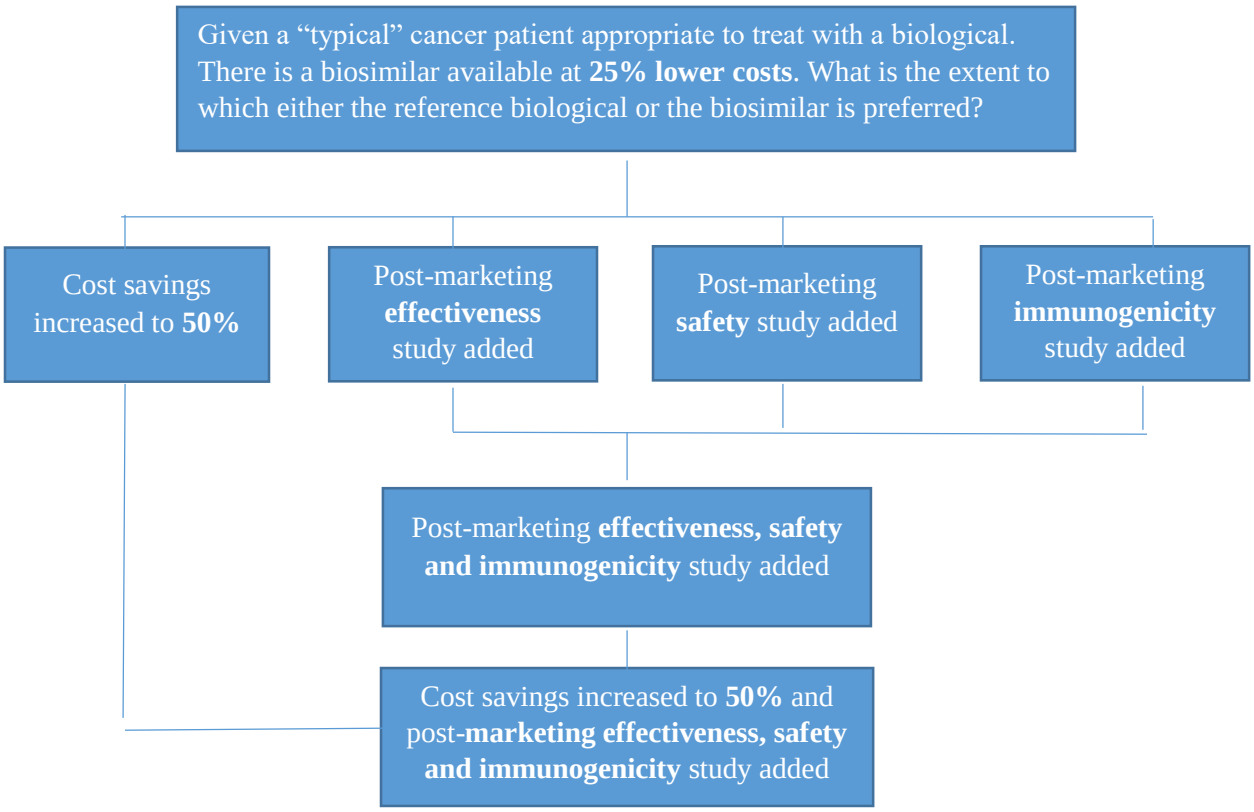


Figure 6 *Combination of MCDA and DCE model*

In the second section of the questionnaire, the respondents were asked to judge a total of six pairwise comparisons, in which the single decision criteria as shown in figure 6 were set out against each other. With reference to the obtained judgments, the pairwise comparison matrix could be drafted. A comparison is made between the weights of the different criteria as derived from these questions. By means of the pairwise comparisons, the relative importance of each of the decision criteria could be obtained. This section of the questionnaire relates to the actual AHP questions, and thus AHP analysis will be used to analyze these data. Thus, it is important to realize that these two sections of the questionnaire differ from each other. The first section primarily focuses upon clarifying the role of cost savings and post-marketing studies, whereas the second section focuses upon determining the priorities and weights of different decision criteria. For that reason, the analysis of the first section differs from the AHP analysis, whereas the analysis of the second section relates to the AHP analysis. The analysis used for the both the preference questions and the pairwise comparison questions will be explained in more detail in paragraph 4.5 (statistical analysis). Finally, in the third and final section of the questionnaire demographic questions were asked. By means of these questions, one was able to perform subgroup analyses. The questionnaire used can be found in Appendix A.

4.5 Statistical analysis

4.5.1 Impact of additional information

As described in the previous paragraph, the first six preference questions can be seen as a combination of AHP questions and DCEs. The 1-9 reciprocal, ordinal scale of the AHP is used to obtain answers regarding the preferences of additional data. This implies that the respondents had 17 possibilities to

answer the questions. The answers are, for that reason, located on a scale from 1 to 17. Therefore, one could state that the data obtained can be seen as continuous data. On this scale, a score from 1-8 represents a preference for the reference biological. A score of 9 implies that the respondent has an equal preference for both products. A score of 10-17 represents a preference for the biosimilar. It is important to note that this response scale is not linear. This is graphically shown by figure 7 below, which indicates that the statistical distance between the answer 1 and 2, for instance, differs from the distance of answer 8 to 9.

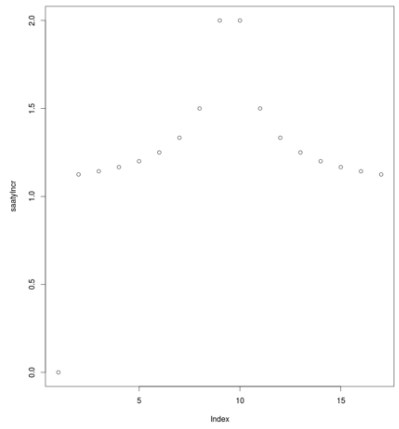


Figure 7 *Response scale AHP*

To investigate if the decision criteria have a positive effect on preferring the biosimilar over its reference biological, R software is used to establish an additive model to perform nonparametric regression. Respondents with a baseline preference that very strongly or extremely preferred the biosimilar were excluded from the analysis. If a respondent had such a baseline preference, it is not possible to investigate the role of adding post-marketing studies for these respondents. All additional information would most likely result in a very strong or extreme preference for the biosimilar. Including these responses would therefore result in a distorted picture of the role of additional information.

The null hypothesis for this model was that extra information on additional cost savings and post-marketing data has no effect on preferring biosimilars over its reference biological:

H0: Q(4-9)-Q3 = 0

H1: Q(4-9)-Q3 ≠ 0

The additive model used five different variables:

Variable	Meaning	Code
X1	Approval data	{1,0,0,0,0}
X2	Post-marketing effectiveness data	{0,1,0,0,0}
X3	Post-marketing safety data	{0,0,1,0,0}
X4	Post marketing immunogenicity data	{0,0,0,1,0}
X5	Additional cost savings	{0,0,0,0,1}

Table 6 *Additive model*

According to these variables, the following codes were established that represent the preference questions asked in the questionnaire:

Question	Code for additive model
Q3	{1,0,0,0,0}
Q4	{1,0,0,0,1}
Q5	{1,1,0,0,0}
Q6	{1,0,1,0,0}
Q7	{1,0,0,1,0)
Q8	{1,1,1,1,0}
Q9	{1,1,1,1,1}

Table 7 *Codes used for additive model*

By means of this additive model, it is calculated if there is a difference between the baseline response to question 3 and the response to the other questions. This model made it possible to analyze if adding information contributed to an increase of the baseline preference. Thus, by means of this model one was able to investigate which decision criteria are preferred by the respondents.

With reference to this nonparametric regression, it was calculated if adding information has an effect on the preferences of the respondents. Additional regressions are performed to investigate if this potential effect is important (significant) or not. It was aimed to model the responses to the preference questions to see if there is an effect on preferring biosimilars, if additional information on biosimilars is provided to that individual. As stated before, a score of 10-17 implies that the respondent favors the biosimilar. When one takes this into account, two groups can be identified for every question: a group that doesn't favor the biosimilar (score of 1-9), and a group that favors the biosimilar (10-17) when different additional data is provided to the individuals. This implies that one gets dichotomous outcomes variables. For that reason, a logistic regression model is fitted. The null hypothesis that has been respected, is that providing additional information will lead to judgments that favor biosimilars. The outcome of this logistic regression model is 'preferring the biosimilar over the reference biologic'. This outcome is defined as having a score response of equal to or greater than 10 on the 1-9 reciprocal, ordinal scale. The fitted logistic regression model used can be written down as:

$$Y(Prob(response>10))=B0 + B1\ (amount\ of\ information\ provided)$$

The amount of information provided, or B1, is as following:

Question	Amount of information provided
Q3	Information on approval data only
Q4	Information on approval data and additional costs savings
Q5	Information on approval data and additional post-marketing effectiveness data
Q6	Information on approval data and additional post-marketing safety data
Q7	Information on approval data and additional post-marketing immunogenicity data
Q8	Information on approval data and additional post-marketing effectiveness, safety and immunogenicity data
Q9	Information on approval data and additional post-marketing effectiveness, safety, immunogenicity data and additional cost savings

Table 8 *Amount of information provided*

The predictor ‘additional information’ is used as a non-linear categorical variable and Q3 is the reference or baseline variable.

With reference to this model, one was able to investigate what additional information and thus which decision criteria are preferred by the respondents. Besides, one was able to investigate what additional information can be associated with higher odds of favoring biosimilars over its reference biologic.

4.5.2 Pairwise comparisons

In the second section of the questionnaire, pairwise comparison questions are asked. The AHP analysis is used to statistically analyze this data. The obtained data is recorded and stored in LimeSurvey. Microsoft Excel version 2013 and the AHP Excel Template Version 2015-06-07[93] are used to analyze the obtained data. First, a consistency test is performed to determine the consistency of the obtained judgements. If the consistency ratio was equal to or higher than 0.20, the judgments were revised or excluded from the analysis. According to the included obtained judgments, the pairwise comparison matrix was drafted of both the group and each individual. With reference to these matrices, the relative weights and the overall priorities were calculated. The overall priority was calculated by means of the normalized principal Eigenvector. Since multiple judgments are obtained, the percentage of consensus between the judgments of the respondents was calculated. This percentage of consensus is calculated with the AHP Excel Template Version 2015-06-07[93]. In this template, Shannon alpha and beta entropy are used to calculate the consensus indicator, which ranges from no consensus (0%) to full consensus (100%)[93]. The consensus indicator gives an indication to what extent the judgments of the respondents are comparable.

4.5.3 Subgroup analysis

Subgroup analyses are conducted by means of the judgments as derived from the pairwise comparison questions. The analysis are performed on the basis of the (individual) geometric means, or priorities, per decision criteria. Individuals were divided into subgroups, of which the relative weights and overall priorities were calculated. In addition, the minimum, maximum and median value per decision criteria were calculated, as well as the mean and the 95% confidence intervals. This analyses made it possible to identify whether demographics, consistency ratios and different reference biologicals can be associated with potential differences within the various criteria.

4.6 Ethics

Once the survey is completed by the respondent, the data is restored in LimeSurvey. Only members of the research team had direct access to the obtained, raw data. Furthermore, the obtained data is not identifiable to a person to ensure confidentiality and anonymity. The respondent had the possibility to stop filling in the questionnaire at any point.

5. Literature review

The outcomes of the MCDA study are related to biological products of which biosimilars currently do not exist. For that reason, a systematic literature review is conducted with regard to biosimilars, as disclosed in the relevant scientific literature. The objective of this literature review was to determine the current scientific knowledge of biosimilars in oncology, thus of biosimilars that currently exist. On the basis of this information one is able to say if post-marketing data regarding biosimilars in oncology is performed. In addition, it is possible to determine what the main focus of these studies was in terms of effectiveness, safety and immunogenicity, and if so, which post-marketing studies are rarely performed. If applicable, the literature review can support the findings of the MCDA survey if it emerges that the amount of studies regarding the equivalence of biosimilars is small, and the respondents of the MCDA survey indicate that additional post-marketing studies are required.

5.1 Study selection

The systematic literature review is carried out by performing a literature search in the databases PubMed and International Pharmaceutical Abstracts (IPA). This literature search resulted in a total of 189 articles. The articles found were screened on the basis of the title and the abstract, to identify articles that were eligible for the review. Criteria to include articles were: (1) the study design must provide a comparison of biosimilars with reference biologicals, (2) the result section must include data, and (3) the article must focus upon oncology. If articles were excluded, the reasons were: (1) the article was an editorial, (2) it was an in vitro study, (3) it was a descriptive summary or (4) it was a phase I or phase II study. After screening the abstract it appeared that 173 articles were not appropriate for the literature review according to the criteria. For that reason, they were excluded from the literature review. Another two articles were rejected from the review, because of the fact that the full text could not be located. A total of 14 studies were found that focused on biosimilars in oncology and complied with the inclusion criteria. Of these 14 studies the reference list was screened for other relevant studies. In addition, the abstracts of the articles that cited one of these 14 studies were screened. However, no additional studies were found to be relevant for this literature review. The flow diagram of this literature search can be found in Appendix B.

5.2 Characteristics of included studies

Of the 14 studies that were identified for the literature review, 10 were observational studies, two were comprehensive reviews, one was a randomized trial and one was a non-randomized survey. All the included studies were published in English and after 2011. The duration of the included studies varied from 3 months up to 36 months. In total, 3.375 participants were involved. Of these 3.375 participants, 3.353 (99.8 percent) were patients who suffered from various types of cancer. The remaining 22 participants involved were healthy donors.

The 14 studies can be divided into two groups. This division is made on the basis of the focus of the studies. The majority of the studies (N=12) focused upon the biosimilar granulocyte colony-stimulating factor (G-CSF). The other two articles focused upon the biosimilar epoetin alfa (EPO). A summary of the characteristics per focus group can be found in table 9 below.

Focus group	Total number of participants	Total number of studies
EPO	137	2
G-CSF	3238	12

Table 9 *Characteristics per focus group*

In all the included studies, a comparison was conducted regarding biosimilars and their reference biologicals in terms of effectiveness, safety or immunogenicity. If applicable, the occurrence of (serious) adverse events was mentioned. The duration of the time in which the outcome was measured varied among the studies. This time could vary from monthly investigations to investigations with a time duration of one year or more.

5.3 Results

The included studies varied widely with regard to the focus group, study design, participants, sample sizes and outcome measures. For that reason, it was decided to focus on describing the outcomes, the main results and both the positive and negative findings, instead of conducting a meta-analysis. A brief qualitative summary of each of the studies can be found in Appendix C. In table 10 below, the sample size characteristics are shown of the 14 included studies in this literature review.

Total number of participants included	**3.375**
Total number of studies included	14
Minimum sample size	20
Maximum sample size	1.302
Mean sample size (±SD)	241±379
Median sample size	106

Table 10 *Sample size characteristics of the included studies*

5.3.1 Summary of evidence

5.3.1.1 Effectiveness

All the included studies (N=14) reported findings about the effectiveness of the biosimilar, and compared these with findings of the reference biological. The comparability of the effectiveness is investigated by using a variety of outcome measures. These outcome measures were, among others, the mobilization of peripheral blood stem cells (PBSC), the white blood cell count, the occurrence of neutropenia or chemotherapy disturbance and the hemoglobin response rate. All studies demonstrate that the findings regarding these outcome measures were broadly similar and comparable. None of the studies reported results indicating that the biosimilar performed significantly worse with regard to the effectiveness. To clarify the obtained results in terms of effectiveness, the studies that focused upon PBSC mobilization will be shown (N=7). The results of these studies can be found in table 11 below.

Author, year	**Outcome Biosimilar**	**Outcome Reference Biological**	**P-value**
Manko et al., 2014[78]	9.1 (0-23)	9.4 (6-48)	-
Publicover et al., 2013[79]	4.5 (min-max: 0.2-43)	4.4 (min-max: 0.5-56)	0.65
Remenyi et al., 2014 (Study I)[80]	6.33 (min-max: 2-17.4)	No control group used	-
Remenyi et al., 2014 (Study II)[80]	5.2 (min-max: 2.22-57.07)	No control group used	-
Schmitt et al., 2013[81]	4.4 (min-max: 2.0-7.3)	4.2 (min-max: 2.1-7.9)	0.75
Sivgin et al., 2013[82]	13.43 (min-max: 8.15-23.38)	7.64 (min-max: 4.09-13.86)	0.013
Yafour et al., 2013[83]	4.1 (min-max: 0.25-4.84)	2.75 (min-max: 1.22-10.3)	0.86

Table 11 *Efficacy in PBSC mobilization: the number of peripheral blood CD34+ cells (x10^6/kg body weight)*

The value provided in this table reflects the median number of peripheral blood CD34+ cells (10^6/kg body weight). The values between the parentheses show the 25th and 75th percentiles. Below, a figure is provided to show these findings graphically. The results of the studies of Remenyi et al. (2014) are excluded from the graph, due to the fact that these studies did not include a control group.

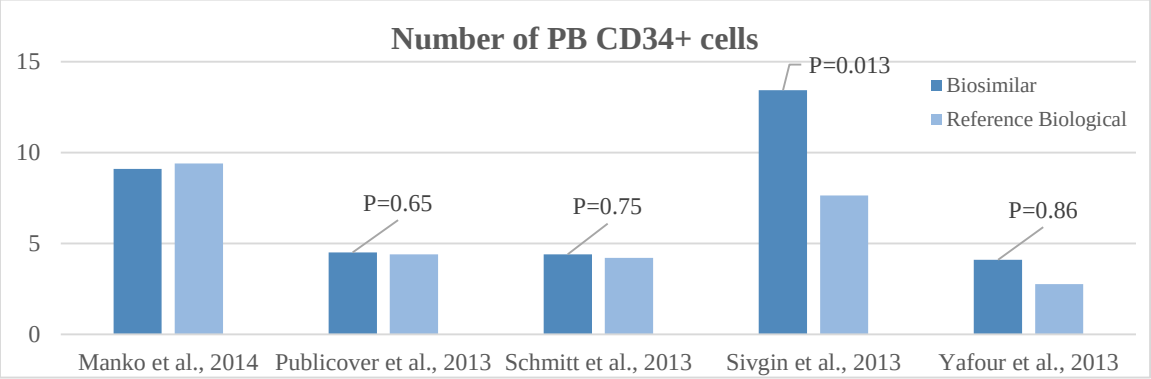


Figure 8 *PBSC mobilization: number of PB CD34+ cells (10^6/kg body weight)*

The findings of the other outcome measures related to the effectiveness can be found in Appendix D.

5.3.1.2 Safety

Of all the included studies, almost 79 percent (N=11) also focused upon the safety of the biosimilar. The frequency and severity of both adverse events and side effects were primarily used to investigate the comparability in terms of safety. The findings of these studies are shown in the following table:

Author, year	(Primary) Outcome Measure(s)	Outcome Biosimilar	Outcome Reference Biological
Gascon et al., 2013[84]	Bone pain	8%	24%
Iannotto et al., 2014[85]	*Lymphoma:*		
	Bone pain	N=3	N=0
	Pneumopathy	N=2	N=2
	Day-100 death	N=1	N=2
	Myeloma:		
	Bone pain	N=6	N=6
	Pneumopathy	N=1	N=0
	Day-100 death	N=0	N=0
Kerkhofs et al., 2012[79]	Unexpected adverse events	None reported	No control group used
Manko et al., 2014[78]	Neutropenic fever	N=11	N=10
	Bone pain	N=17	N=19
Remenyi et al., 2014[80]	Neutropenic fever	N=45 (64%)	No control group used
	Engraftment syndrome	N=9 (13%)	No control group used
	Poor graft function	N=1 (1.4%)	No control group used
	Deaths (treatment related)	N=0	No control group used
Sagara et al., 2013[87]	Incidence of febrile neutropenia	34.6% (N=36)	No control group used
Salesi et al., 2012[88]	Unexpected adverse events	None reported	No control group used
Schmitt et al., 2013[81]	Side effects (arthralgia)	N=6 (54.5% donors)	N=6 (54.5% donors)
Schmitt et al., 2014[89]	Increase in toxicity	None reported	No control group used

Verpoort & Mohler, 2012[90]	Unexpected adverse events	None reported	None reported
Yafour et al., 2013[83]	Bone pain	33,3% (N=3)	20% (N=2)
	Headache	33,3% (N=3)	10% (N=1)

Table 12 *Outcome measures with outcomes related to safety*

5.3.1.3 Immunogenicity

In total, 14 percent of the studies (N=2) focused on both efficacy, safety and immunogenicity. The main outcome measures to measure immunogenicity were the occurrence of antibodies, and the frequency and severity of immune responses. The findings of these two studies with regard to the immunogenicity of the biosimilar can be found in table 13 below.

Author, year	(Primary) Outcome Measure	Outcome Biosimilar	Outcome Reference Biological
Gascon et al., 2013	Production of G-CSF antibodies	No neutralizing antibodies found (N=316)	No control group used
Sagara et al., 2013	Production of G-CSF antibodies	No anti-G-CSF antibodies found before and after treatment with the biosimilar (N=102)	No control group used

Table 13 *Outcome measures with outcomes related to immunogenicity*

5.4 Conclusion

The main focus of the studies included in the literature review was on effectiveness. All included studies demonstrated that the findings regarding these outcome measures were broadly similar and comparable. None of the studies reported results indicating that the biosimilar performed significantly inferior with regard to the effectiveness. The results of the PBSC mobilization show that the differences in the number of PB CD34+ cells between the biosimilar and the reference biological are not statistical significant. Only the study of Sivgin et al. (2013) found a significant difference. This difference is, however, in favor of the biosimilar since the patients who received the biosimilar created more PB CD34+ cells.

According to the findings, no remarkable differences in safety profiles were found between both the reference biological and the biosimilar. If adverse events occurred, the frequency and severity of them were comparable between both groups. With regard to immunogenicity, the biosimilar and the reference biological were found to be comparable in both included studies. The fact that only 14 studies are found to be relevant for this literature review suggest that the number of studies regarding biosimilars in oncology is limited. In addition, the median number of patients included in these studies is relatively small (N=106). This small sample size means that detecting uncommon problems with regard to safety and immunogenicity is unlikely. Besides, the current data that is published provides minimal evidence regarding safety and immunogenicity of biosimilars in the area of oncology. In addition, a large proportion of the surveys did not use a control group in which patients received the reference biological.

6. Results

6.1 Inclusion of respondents

A total of 210 emails were sent to hospitals, health insurance companies, pharmaceutical companies and research centers. A total of 34 respondents completed the questionnaire. The response rate cannot be calculated, since the snowball sampling method was used. Of the 34 respondents, 74% was male (N=24) and 26% was female (N=9). A total of 5 respondents were aged between 31-40 years (15%), 13 were aged between 41-50 years old (38%), 12 between 51-60 years old (35%) and three were 60 years or older (9%). The majority of the respondents were clinical care provider (68%). The remaining 32% (N=12) represent a variety of primary activities. A total of 10% were researcher, 7% were teaching at a university or medical facility, 7% were employees of a pharmaceutical industry, 3% were policy maker, 3% were involved in the benefit structure of a health insurance company, 3% were working in the biotech industry and 3% were patient representatives. A total of 22 respondents (65%) indicated that they had ever prescribed, dispensed or administered a biological product. A total of three respondents (10%) answered that they had received a biological product as a patient. Of these three respondents, one mentioned that the respondent or a close family member had experienced a serious or a life-threatening adverse reaction from a drug or biological product. The average baseline preference observed without exclusion is 10.3. The baseline preference of this respondent was represented by a score of 6, which equals a score between moderately preferring and strongly preferring the reference biological product.

6.2 Preferences of additional information

Of the 34 respondents, 7 had a baseline preference that very strongly or extremely favored the biosimilar. These 7 respondents (21%) were excluded from the analysis of the preference questions. Of the 27 respondents included, the mean baseline preference was 8.85. Of these 27 respondents, 20 had ever prescribed, dispensed or administered a biological product. Their mean baseline preference was 9. The group who never prescribed, dispensed or administered a biological product (N=5) had a mean baseline preference of 8.6. Of two respondents it was unknown if they had ever taken any of these actions.

In the table below, the results are provided as obtained from the nonparametric regression in the additive model.

Coefficients	Estimate	Std. Error	Z value	Pr(>|z|)
(Intercept)	0.2581	0.2547	1.013	0.312207
X1	NA	NA	NA	NA
X2	2.1924	0.3762	5.828	2.46e-08
X3	1.3405	0.3762	3.564	0.000466
X4	1.2294	0.3762	3.268	0.001291
X5	1.0645	0.3601	2.956	0.003526

Table 14 *Additive model regarding the role of additional information corrected for baseline preferences*

The figure on the next page is a graphical representation of the results of the additive model. It shows how the preferences change from the baseline preferences, when additional (post-marketing) data is provided to the respondent. The height of the bar represents a higher change in preference, when that particular information is provided to the respondent.

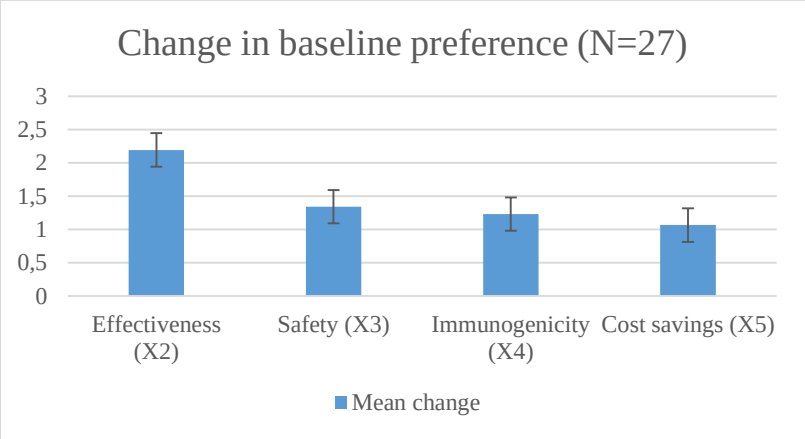


Figure 9 *Change in baseline preference after providing additional data*

The results of this crude logistic regression are provided below in table 15 (N=27).

Coefficients	Estimate	Std. Error	Z value	Pr(>|z|)
(Intercept)	0.07411	0.38516	0.192	0.84742
Q(4)	0.61904	0.56126	1.103	0.27006
Q(5)	1.17865	0.60216	1.957	0.05032
Q(6)	0.61904	0.56126	1.103	0.27006
Q(7)	0.45652	0.55423	0.824	0.41011
Q(8)	2.45162	0.82967	2.955	0.00313
Q(9)	2.45162	0.82967	2.955	0.00313

Table 15 *Regression table regarding the role of additional information*

6.3 Priorities of decision criteria

Of the 34 respondents, 23 (68%) met the consistency ratio threshold (>0.20). These 23 respondents were included for the analysis of the criteria weighting questions. With reference to the judgments of the respondents, the pairwise comparison table was established. The findings are presented in table 16 below.

	Effectiveness	Safety	Immunogenicity	Cost Savings
Effectiveness	1.00	0.88	1.83	3.00
Safety	1.17	1.00	1.96	3.44
Immunogenicity	0.69	0.56	1.00	2.23
Cost savings	0.38	0.29	0.47	1.00

Table 16 *Pairwise comparison table findings*

The weights and the overall priorities as derived from the judgments of the respondents (N=23) are showed in table 17. The AHP group consensus was 77.4%.

	Effectiveness	Safety	Immunogenicity	Cost Savings	Priorities (%)
Effectiveness	0.31	0.32	0.35	0.31	32.3
Safety	0.36	0.37	0.37	0.36	37.0
Immunogenicity	0.21	0.20	0.19	0.23	20.4
Cost savings	0.12	0.11	0.09	0.10	10.4

Table 17 *Pairwise comparison matrix*

6.4 Subgroup analyses

6.4.1 Consistency ratio

The first subgroup analysis that was performed concerns the consistency ratio. The respondents are divided into two groups. The division is made on the basis of the consistency of their judgments. One group regards a consistency ratio of 0-10%. The second group consists of judgments with a consistency ratio of 10-20%. The subgroup analysis is performed with regard to the (individual) priorities per decision criteria. The overall priorities as derived from the judgments of the respondents are showed in the table below.

Group (n)	Effectiveness (%)	Safety (%)	Immunogenicity (%)	Cost savings (%)
Consistency 0-10% (N=16)	33.9	35.1	21.5	9.4
Consistency 10-20% (N=6)	27.4	42.1	17.2	13.3

Table 18 *Overall priorities per consistency ratio group*

The consensus in the group with a consistency ratio of 0-10% was 81.5%. The consensus among the members of the group with a consistency ratio of 10-20% was 60.6%. The figure below shows the differences in priorities among both subgroups identified.

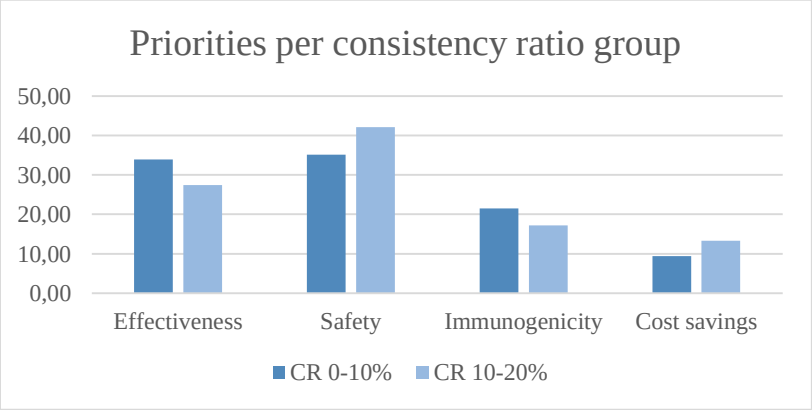


Figure 10 *Subgroup analysis: differences in priorities per consistency ratio group*

In addition, subgroup analysis was performed regarding the answers about the baseline preference of biosimilars in oncology. The mean preference of the group with a consistency score between 0-10% was 10±1.50 (95 % CI 8.50-11.50). The mean preference of the group with a consistency score between 10-20% was 9.33±4.19 (95 % CI 5.14-13.53). The minimum, median and maximum values of the priorities of both groups for each of the decision criteria are shown by box plots in the figure on the next page.

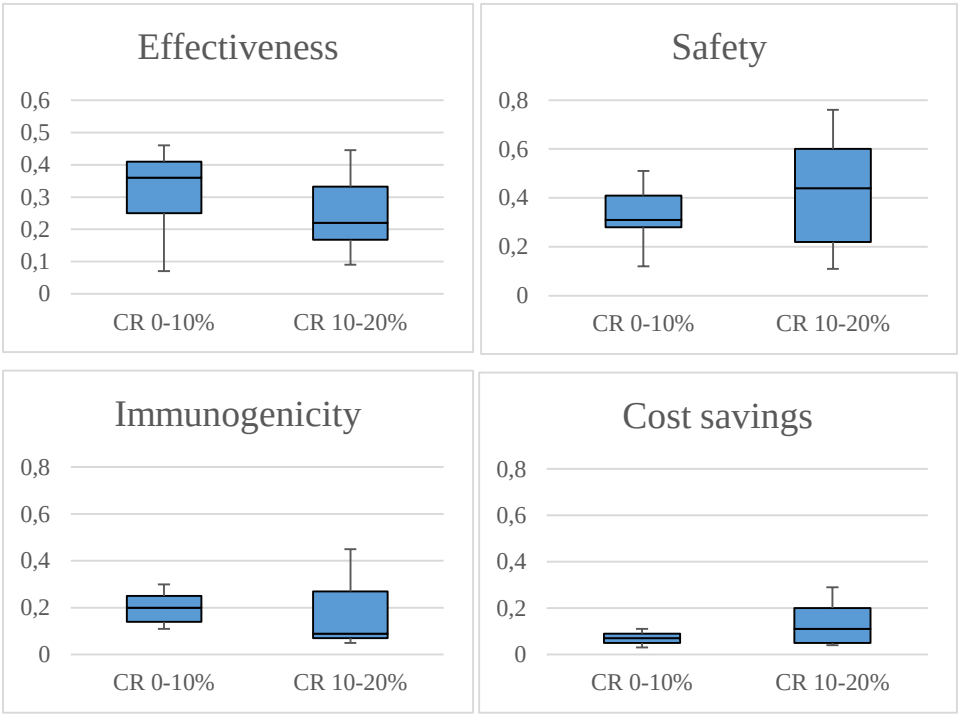


Figure 11 *Box plots of the priorities per criteria per consistency ratio*

6.4.2 Chosen biological product

Since the respondents were asked to choose between the biological product that they are most familiar with, a subgroup analysis is performed with regard to the chosen biological product. A total of four groups can be distinguished, namely (1) CD20 antigen inhibitor (N=9), (2) HER-2 inhibitor (N=8), (3) human vascular endothelial growth factor (N=4) and (4) tyrosine kinase inhibitor (N=2). The biological product epidermal growth factor receptor inhibitor was not selected by any respondent. The overall priorities (normalized principal Eigenvector) per decision criteria for every subgroup are showed in the table below.

Group (n)	Effectiveness (%)	Safety (%)	Immunogenicity (%)	Cost savings (%)
CD20 antigen inhibitor (N=9)	32.6	39.0	16.2	12.1
HER-2 inhibitor (N=7)	32.2	32.0	26.6	9.2
Human VEGF inhibitor (N=4)	25.9	45.5	17.6	11.0
Tyrosine kinase inhibitor (N=2)	41.3	29.6	23.1	6.0

Table 19 *Overall priorities per chosen biological product*

The consensus in the CD20 antigen group was 72.3%. The consensus among the members of the group who had chosen for HER-2 inhibitor was 77.1%. The consensus among the provided judgments of the respondents of the human VEGF and the tyrosine kinase inhibitor were 77.5% and 90.1% respectively. The figure below shows the differences in priorities among these four identified subgroups.

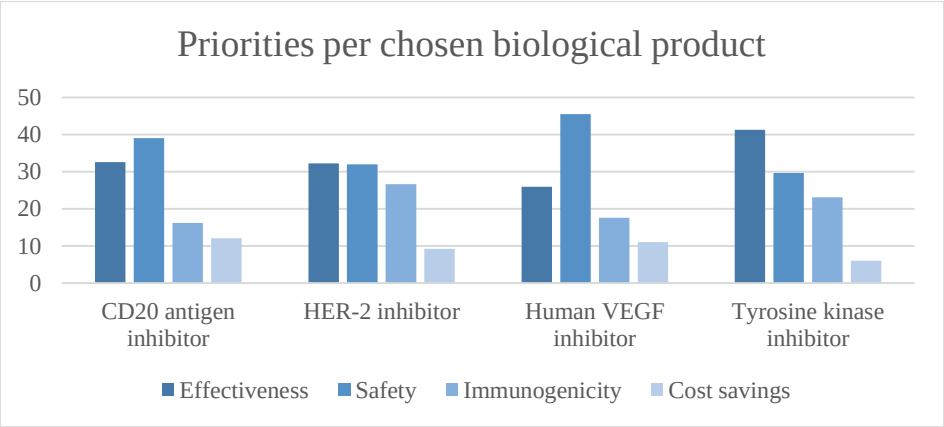


Figure 8 *Subgroup analysis: differences in priorities per chosen biological product*

For these four different groups, a subgroup analysis is performed with regard to the answers about the baseline preferences of biosimilars. The mean preference of the CD20 antigen group was 9.11±2.82 (95 % CI 6.29-11.93). The mean preference of the group that chose for the biological product HER-2 inhibitor was 10.00±2.75 (95 % CI 7.25-12.75). The group that chose for human vascular endothelial growth factor had a mean baseline preference of 10.75±3.14 (95 % CI 7.61-13.89). Finally, the mean baseline preference of the tyrosine kinase inhibitor group was 10.50±1.34 (95 % CI 9.16-11.84). Figure 13 below shows the box plots of each of the subgroups per decision criteria.

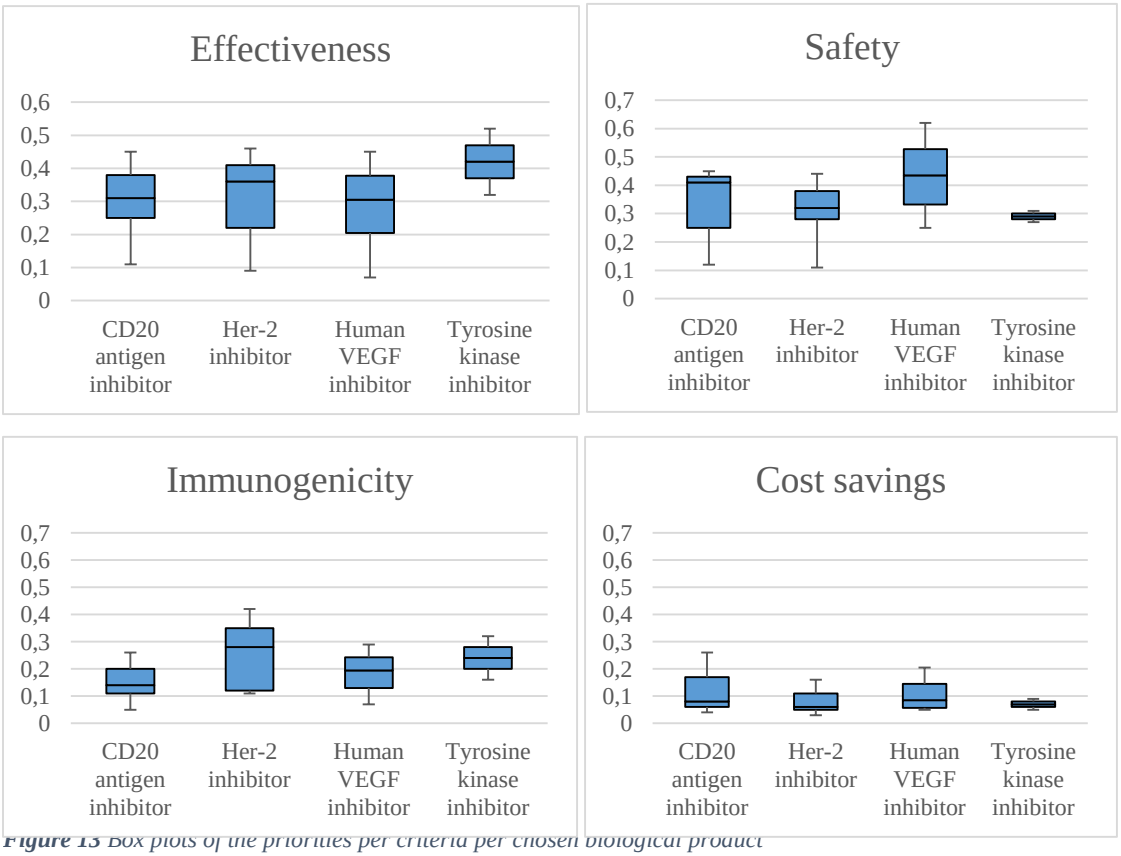


Figure 13 *Box plots of the priorities per criteria per chosen biological product*

6.4.3 Background

Multiple different stakeholder groups completed the online questionnaire. For that reason, a subgroup analysis is performed regarding the background of the respondents. The background represents the primary activity of each of the respondents. A total of four groups are distinguished, which are (1) physicians (N=4), (2) oncologists (N=7), (3) hematologists (N=4), and (4) other primary activities (N=8). The subgroup that is encompassed by the term 'other' represents a variety of primary activities. The primary activities represented in this group, are a patient representative (N=1), a pharmacist (N=1), a clinical pharmacologist (N=1), an EMEA biosimilar market access expert (N=1), a laboratory technician (N=1), a hospital pharmacist (N=1), a clinical biological researcher (N=1) and a researcher (N=1). The overall priorities per decision criteria for each of the four subgroups can be found in table 20 below.

Group (n)	Effectiveness (%)	Safety (%)	Immunogenicity (%)	Cost savings (%)
Physician (N=4)	21.2	56.7	11.3	10.8
Oncologist (N=7)	40.1	31.2	18.3	10.4
Hematologist (N=4)	35.4	36.7	19.3	8.6
Other (N=8)	29.0	32.3	28.5	10.2

Table 20 *Overall priorities per background group*

The consensus among the judgments of the physicians was 83.6%. For the oncologists and the hematologists, the consensus was 85.0% and 85.8% respectively. The final subgroup, encompassed by the term 'other', had a consensus of 63.8%. Figure 14 provides a graphical image of the overall priorities per primary activity:

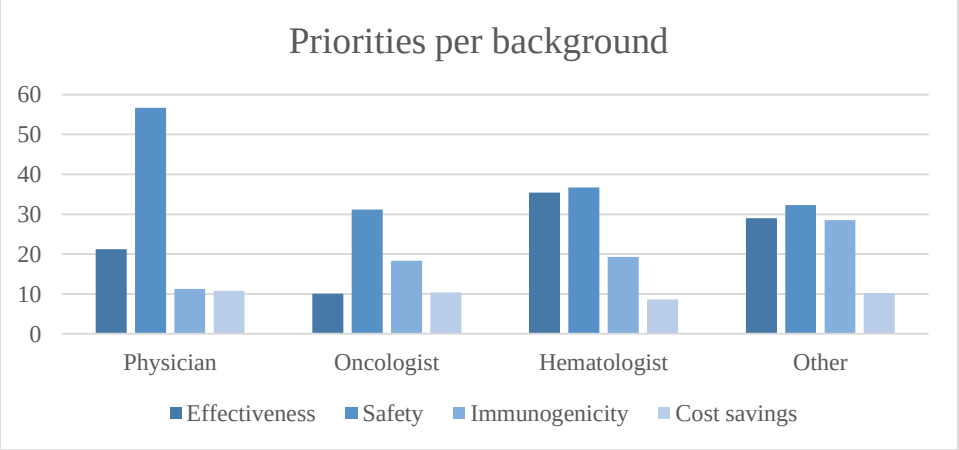


Figure 14 *Subgroup analysis: differences in priorities per background*

The mean baseline preference of the pharmacists can be represented by a score of 9.75±4.89 (95 % CI 4.86-14.64). The judgments of the oncologists imply that their baseline preference has a mean score of 9.14±2.07 (95 % CI 7.07-11.21). The mean preference of the hematologists was 8.00±6.33 (95 % CI 1.67-14.33). All the remaining judgments combined into the final, 'other', group had a mean baseline preference of 11.38±6.87 (95 % CI 4.51-18.24). Furthermore, box plots are drafted to show the differences in minimum, median, and maximum priorities according to the different backgrounds.

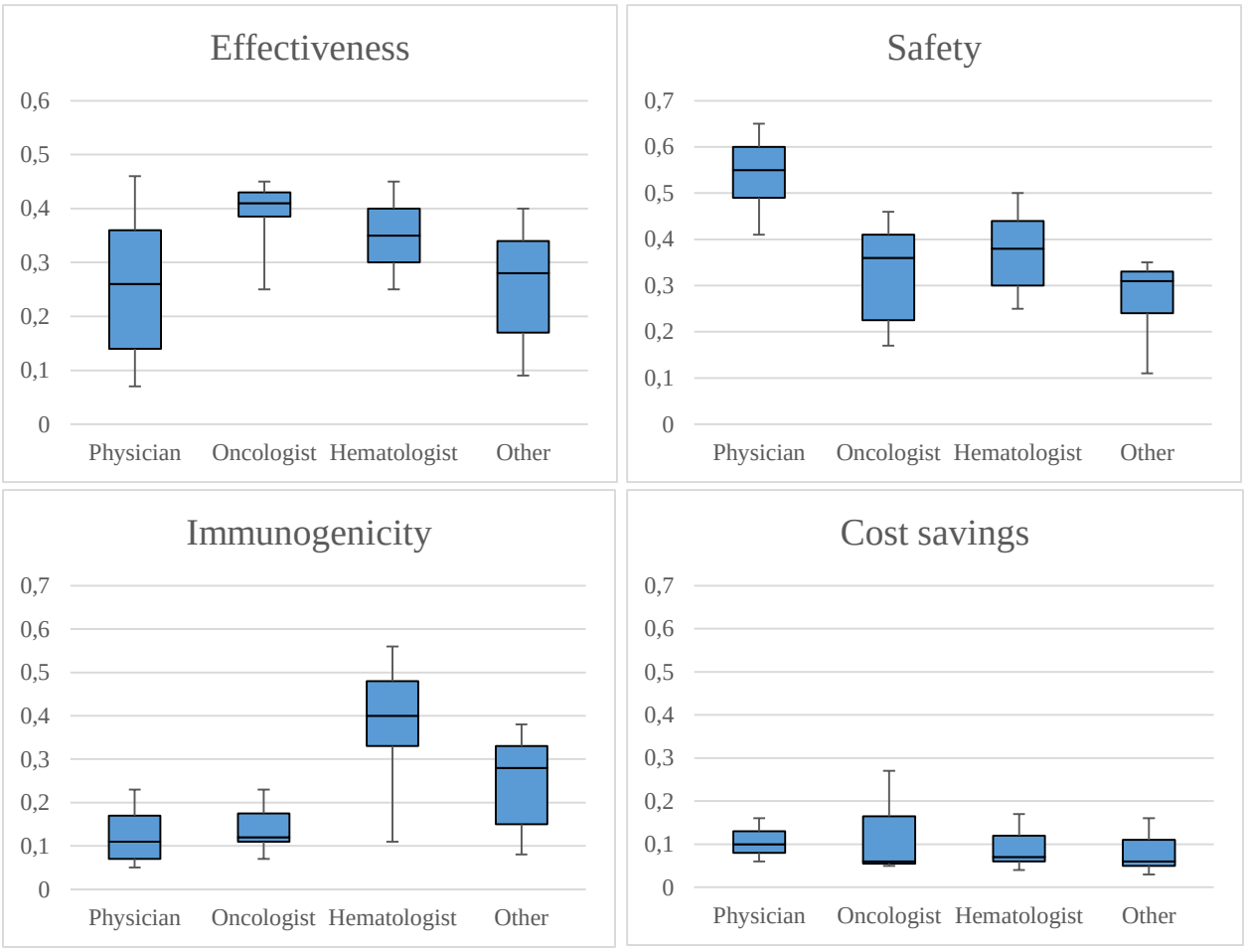


Figure 15 *Box plots of the priorities per criteria per background*

6.5 Other criteria

Respondents were asked, if applicable, to list any other important criteria in the decision whether to use a biosimilar instead of its reference biological product. Even though one respondent mentioned that effectiveness, safety and immunogenicity are most important, other criteria are mentioned that are also relevant to the respondents in the decision making process. Outcome registration within the profession, adaptation by scientific associations and patient friendliness are considered to be important. In addition, a structured follow-up of the patient is required by the respondents. According to one respondent, the physical-chemical quality standards as described for investigational medicinal products (IMP) are of great importance, as well as the substances that are added in addition to the active substances. Respondents, however, indicated that they are not willing to switch to a biosimilar when the patient already receives its reference biological product and this product shows good (clinical) results. It was also mentioned by respondents that new patients have less objection towards new medical products in comparison with existing patients who have to switch from (oncology) drugs.

7. Conclusion

With this research, the criteria that stakeholders in oncology consider to be decisive with regard to the uptake of biosimilars in the Netherlands are studied. Furthermore, this research aimed to elucidate the interest of post-marketing studies. First, the results per sub question will be described. With reference to these results, an overall conclusion regarding the research question of this dissertation will be presented.

7.1 Conclusion per sub question

7.1.1 Conclusion sub question 1

From the literature review it can be concluded that the scientific knowledge of biosimilars in oncology is limited. It is found that the scientific knowledge of biosimilars in oncology focused upon EPO and G-CSF, of which biosimilar products are currently approved. Post-marketing studies are rarely performed, and if they are performed their main focus is on effectiveness. In addition, the median number of patients included in these studies is small. The post-marketing studies that are performed show no remarkable differences between biological products and their biosimilar with regard to effectiveness, safety and immunogenicity profiles. More post-marketing studies are, however, required before one is able to conclude with more certainty that biological products and their biosimilars are equivalent in terms of effectiveness, safety and immunogenicity.

7.1.2 Conclusion sub question 2

From the results of the additive model it can be observed that additional information contributes to an increase of the preference, baseline score. This increase in preference score applies for both additional cost savings data as well as additional post-marketing data on effectiveness, safety and immunogenicity. From these results it can be concluded that adding information has a positive effect on the preferences of the respondents. From the results of the logistic regression it can be observed that providing all post-marketing data along with approval data, and all post-marketing data along with approval data as well as additional cost savings, is associated with significantly higher odds ($p<0.01$) of preferring biosimilars over reference biologics. It is also observed that post-marketing effectiveness data along with approval data was associated with significantly higher odds ($p=0.05$) of favoring biosimilars over biologics. According to the findings, post-marketing studies that provide safety or immunogenicity data only, as well as additional cost savings, did not result in significantly higher odds of preferring the biosimilar over its reference biological. In contrast, when the four decision criteria are compared with each other, post-marketing safety data is considered to have the highest relative importance to the respondents with regard to the uptake of biosimilars (weight = 0.37). After post-marketing safety data, post-marketing effectiveness data was considered to be the most important (weight = 0.323), followed by post-marketing immunogenicity data (weight = 0.204) and cost savings (weight = 0.104).

7.1.4 Conclusion sub question 3

From the results, it is observed that the preferences vary when different potential biosimilars are taken into account by the respondents. Of the five most prescribed oncology drugs according to sales, the respondents in this research mostly prescribed or are most familiar with CD20 antigen inhibitor and HER-2 inhibitor. The respondents in the groups CD20 antigen inhibitor and human VEGF had the highest preference for safety, followed by effectiveness, immunogenicity and cost savings. The respondents in the HER-2 inhibitor and tyrosine kinase inhibitor groups, on the other hand, preferred effectiveness the most, followed by safety, immunogenicity and cost savings. With regard to the cost savings, the priorities of all

four groups are in accordance. The other three decision criteria are subjected to more variability. The highest variability is observed between safety and effectiveness. From the findings it can be concluded that the respondents that chose for human VEGF had the highest overall preference for safety. The respondents that chose for the reference biological tyrosine kinase inhibitor had the highest overall preference for effectiveness. Furthermore, the respondents that chose HER-2 inhibitor had nearly similar preferences for safety and effectiveness. These findings show differences between the judgments of the respondents. It was, however, not possible to calculate significant differences because of the small sample sizes of the distinguished groups.

In addition, the mean baseline preference differed between the four groups identified. The respondents of the CD20 antigen inhibitor and the human VEGF slightly groups preferred the biosimilar over its reference biological product. The respondents of the HER-2 inhibitor and the tyrosine kinase inhibitor had a higher mean baseline preference and almost moderately preferred the biosimilar. The consensus of the judgments within the groups is relatively high. The 95% confidence intervals of all four groups is, however, broad.

7.1.5 Conclusion sub question 4

From the results of the subgroup analysis of the various stakeholder groups included, it can be concluded that there are differences in preferences with regard to the uptake of biosimilars in oncology. Physicians highly preferred safety, followed by effectiveness, immunogenicity and cost savings. The same ranking is observed in the judgments of the hematologists and the group encompassed by the term ‘other’. The judgments of the last two groups have, however, less remarkable differences in preferences between safety and the other three decision criteria. In contrast, oncologists highly prefer effectiveness, followed by safety, immunogenicity and cost savings. With reference to the boxplots, it is observed that the judgments regarding the costs savings are broadly similar between the groups. More variability in the judgments of the other decision criteria are found. Pharmacists have highly variable preferences regarding effectiveness, whereas oncologists have highly variable preferences regarding safety. In addition, it can be concluded that there is more variability in the preferences regarding safety studies when compared to the preferences of immunogenicity studies.

The mean baseline preferences also differ per stakeholder group. The pharmacists and oncologists slightly prefer biosimilars over its reference biological. Hematologists prefer the reference biological over the biosimilar, with a score that represents a preference between equally prefer and moderately prefer the biological. The judgments combined into the ‘other’ group show that these respondents have a moderate to strong preference for the biosimilar. The consensus among the judgments of physicians, oncologists and hematologists is relatively high. The consensus among the judgments of the other respondents is, in contrast, relatively low. Furthermore, all groups have a broad 95% confidence interval.

7.1.6 Conclusion sub question 5

When the preferences of the pilot study are compared with the preferences of this research, it can be concluded that they differ. According to the pairwise comparison matrix of the pilot study, post-marketing immunogenicity studies are considered to have the highest preference, followed by post-marketing studies regarding effectiveness, safety and cost savings. From the pairwise comparison matrix of the current study, it can be concluded that post-marketing safety studies have the highest preference, followed by post-marketing effectiveness, immunogenicity and cost savings. In addition, the range of the priorities in the pilot study is relatively small. From the current study, it can be observed that the range is relatively

large. Thus, larger differences in preferences are found between the observed decision criteria when compared with the pilot study.

From both the studies it can be concluded that preferences differ among subgroups. Furthermore, the results of both studies show that additional cost savings did not result in having a high preference by any subgroup, and thus don't play a large role in the decision-making regarding biosimilars in oncology.

7.2 Conclusion research question

With reference to both the literature review and the survey, it can be concluded that the performance of post-marketing studies is required which are able to demonstrate the comparability of biosimilars with their reference biological product. The current post-marketing experience with biosimilars in oncology is limited, and this survey demonstrates that the expected cost savings of biosimilars alone are very unlikely to result in favoring biosimilars over biologic products. Additional post-marketing effectiveness data is found to increase the preference, baseline score towards biosimilars, and is associated with higher odds of favoring biosimilars over its reference biological. The combination of post-marketing data regarding effectiveness, safety and immunogenicity along with the approval data will most likely result in favoring biosimilars over biological products. The combination of all these post-marketing data, along with approval data and additional cost savings has exactly the same preference as the combination without the additional cost savings. This conclusion depends on the assumption that post-marketing studies show that biosimilars are comparable with their reference biological products. It can, however, be concluded that the additional cost savings play a minimal role in the decision-making process regarding the uptake of biosimilars, and additional post-marketing studies are required.

The ranking of the decision-criteria varied between different chosen reference biological products and different stakeholder groups. The most important decision criteria are, however, for every respondent the post-marketing safety and the post-marketing effectiveness data. This is in accordance with the pairwise comparisons, where the respondents indicated that post-marketing safety data and post-marketing effectiveness data have the highest relative importance.

In conclusion, stakeholders in oncology prefer post-marketing effectiveness data and post-marketing safety data along with approval data in the case of biosimilars in oncology. The combination of all post-marketing data along with approval data will most likely result in preferring biosimilars over its reference biological. Post-marketing studies will play a major role in the potential uptake of biosimilars in oncology in the Netherlands, and are required before their implementation on a large scale can be realized.

8. Discussion

This study shows that the AHP method was valuable in obtaining preferences of individual stakeholders in oncology regarding biosimilars. Before a significant uptake of biosimilars in oncology can be expected in the Netherlands and in Europe, it is necessary to take the preferences of stakeholders into account. The results of this research show that the potential cost savings of biosimilars play a small role in the decision to prescribe them. Post-marketing immunogenicity studies, but most of all post-marketing effectiveness and post-marketing safety studies could, in contrast, play an important role in their uptake. One should, however, be careful with the interpretation of the results of this study. The sample size of this study is small. This small sample size is partly a result of the relatively high exclusion rate, due to both inconsistent judgments and very strong or extreme baseline preferences for biosimilars. Excluding these judgments from the analysis could have had an impact on the validity of the obtained outcomes. In the analysis of the results, judgments with a consistency ratio lower than 0.20 were included. Saaty, however, states that judgements are consistent when their consistency ratio is lower than 0.10. The subgroup analysis of this study shows that the ranking is similar when a consistency ratio threshold of 0.20 is maintained. Therefore, it could be questioned if this threshold of 0.10 should be retained. If this threshold was maintained, valuable judgments would have been lost, and the sample size would have been smaller. It should, however, be noted that the group with a consistency ratio of 10-20% had higher variability in their judgments. A reason for this finding might be that the consensus in the group with a consistency ratio of 0-10% was almost 20% higher in comparison with the group with a consistency ratio of 10-20%.

With regard to the small sample size, it should be noted that representable and reliable results are obtained with the AHP when applied in both (small) representative panels and large groups. The fact that questionnaires are used implies, however, that subjective judgments are obtained. The subjective judgments provide an overview of the preferences of the stakeholders, but they don't provide statistical representativeness. In addition, statistical analyses are performed to investigate if the decision criteria have an effect on the preferences of the respondents, and if this effect is important (significant) or not. According to these analyses, it was investigated what additional information can be associated with higher odds of preferring biosimilars over its reference biological. This association replies, however, with the baseline score as provided by the respondents, i.e. bias towards either biosimilars or its reference biological could have a possible effect on the results. Further research is required to investigate if the obtained judgments and priorities are generalizable and representable in the Dutch oncology setting.

From the results it can be observed that the standard deviations of the priorities are broad and the level of inconsistency is relatively high. This might be caused by the small sample size of the survey. Another explanation could be that the preferences of the respondents are diverse, partly due to their various backgrounds. Even though the preferences are highly variable, post-marketing effectiveness data and post-marketing safety data are obviously preferred, with cost savings as the least important criteria. This sequence in preferences is also mentioned by the American Society of Clinical Oncology (ASCO)[93] in the oncology setting. Hummel et al. mentioned that many decision criteria and decision alternatives, and consequently a long (online) questionnaire, could potentially result in more inconsistent judgments[75]. This survey used a small number of decision criteria and decision alternatives, resulting in a small number of pairwise comparisons. The level of inconsistency is, however, still relatively high. This can be partly explained by the consistency ratio threshold of 0.20. Lower levels of inconsistency could be obtained when a larger consistency ratio is maintained. Besides, a representative panel could be used to increase the level of consistency.

The AHP method typically uses one, fixed alternative. In this survey, respondents could choose between five alternative biological products. The product selected was used in the description of the case example. Thus, instead of one general case example, the case example differed per chosen biological product. This was necessary, since multiple stakeholders were approached and every respondent had to be familiar with (one of) the product(s) listed. It must be realized that the outcome of the survey depends on the reference biological that is being evaluated by the respondent. The fact that different reference biologicals are evaluated, in combination with the small sample size, results in the fact that it is hard to generalize the obtained findings. It is, however, important to have an idea of which reference biological product the respondents are considering when completing the online questionnaire. Every biological product differs in effectiveness, potential adverse events and duration of action. These differences are reflected by the obtained results of this survey. The respondents who choose for the product tyrosine kinase inhibitor, e.g. Imantinib, had the highest preference for post-marketing effectiveness studies. Imantinib plays a major role in stopping or diminishing tumor growth and progression and is generally used to treat patients with leukemia, for instance chronic myeloid leukemia (CML). Current effectiveness of Imantinib is solely based upon studies with newly diagnosed CML patients[95]. Besides, it is known that Imantinib could decrease the drugs' blood levels, resulting in decreasing the effectiveness of this particular cancer treatment[96]. These might be reasons that respondents, who are most familiar with this tyrosine kinase inhibitor, prefer post-marketing effectiveness data of biosimilars of this reference biological.

The respondents who chose human vascular endothelial growth factor inhibitor, e.g. Bevacizumab, had the highest preference for post-marketing safety studies. Bevacizumab is used to attack the protein that allows the cancer cells to grow blood vessels[97]. Concerns are that the usage of Bevacizumab might interfere with the processes in the body that regulate the growth of blood vessels. Consequently, adverse events as bowel perforation, coronary artery disease and peripheral artery disease might worsen the health condition of the patient[98,99]. In addition, bevacizumab is an off-label drug. In 2012, a head-to-head trials are performed with regard to Bevacizumab and the golden standard, licensed, Ranibizumab. From this study, it could be concluded that the usage of Bevacizumab resulted in higher risks of adverse events[100]. These might be reasons that the respondents who chose human VEGF prefer post-marketing safety studies before they are willing to prescribe its biosimilar product. It should, however, be noted that in the case of tyrosine kinase inhibitor as well as in the case of human VEGF only two respondents completed the questionnaire. Further research is required to obtain a sample size that is able to generalize the findings.

From the subgroup analysis regarding the respondents' background it was found that preferences differ per stakeholder group. First, hematologists moderately prefer the reference biological over the biosimilar, where oncologists tend to have a more equal preference between both products. When biosimilars in oncology are compared with biosimilars in other application fields, it is important to note that the characteristics of cancer patients differ from other patients. In the field of oncology, drugs are mostly prescribed to patients whose immune system is suppressed. This indicates that these patients have an increased risk of complications. In addition, drugs are also prescribed to healthy donors, who don't have any therapeutic benefit of receiving the particular drug[101]. For this reason, oncologists and hematologists are required to have a certain amount of knowledge about the equivalence of biosimilars. The current knowledge about biosimilars in oncology is, however, limited, which might explain their skeptical attitude towards biosimilars. In contrast, the physicians and the other backgrounds' as encompassed by the term 'other' have a baseline preference which slightly to moderately favors the biosimilar over its

reference biological. There is a possibility that the respondents in these groups have more knowledge about biosimilars from other application fields. Currently there are, however, no studies that mention similar findings.

In addition, oncologists prefer post-marketing effectiveness studies over other post-marketing data. Even though biological products are said to be more targeted when compared to other drug treatments, healthy cells are still attacked and killed by the cancer drug. Consequently, the immune system of the patient is impaired and suppressed. As a result, the patient might experience toxic side effects. Oncologists, however, consider a certain benefit-risk ratio in treating their cancer patients, where some side effects are tolerated if the benefits of the drug outweigh the risks. Cancer is a life-threatening disease, and the effects of cancer on the human body can be more serious than the side effects of the treatment[102]. Therefore, in the benefit risk evaluation of the treatment, it can be stated that oncologists relatively prefer the benefit of an effective treatment over its potential risks, explaining their preference for post-marketing effectiveness data regarding biosimilars over the other post-marketing data.

In contrast with the preferences of the oncologist, the physicians prefer post-marketing safety data over the other post-marketing data. Besides, physicians have highly variable judgments with regard to post-marketing effectiveness data. Reasons for these findings might be the small sample size of this stakeholder group or that this is a representation of the individual preferences of the physicians. Second, since a distinction is made between oncologists and physicians, chances are that, partly due to the snowball sampling method, the questionnaire is forwarded to physicians of broad specialties. Therefore, it is possible that the physicians included in this survey could have different preferences regarding biosimilars due to their backgrounds. The subgroup analyses showed interesting findings with regard to different preferences among the various subgroups. These differences are showed by graphs and box plots. When one takes a look at these figures it stands out that, in some cases, the median varies widely from the average values. A reason might be the level of variability in judgments among the various stakeholder group, in combination with the small sample sizes of some of the stakeholder groups included in the subgroup analyses.

According to the findings, the level of variability in judgments regarding post-marketing safety data is higher than the level of variability in judgments about post-marketing immunogenicity data. Even though definitions of both decision criteria were provided in the questionnaire, it is possible that the term immunogenicity is more specific and clear to the respondent than the term safety. Safety can relate to relatively mild adverse events, but also relate to life-threatening adverse events. Notwithstanding a distinction is made between immunogenicity and safety, it is a possibility that the respondents see immunogenicity as a component of safety. Thus, the respondents might have evaluated safety as a broader concept, resulting in being more concerned about safety than about immunogenicity. Although the obtained findings are variable, all judgments regarding the potential cost savings of biosimilars in this survey are highly similar with less variability. This might be explained by the fact that it is known that biosimilars have the potential to lower costs, but this criteria is not of major importance to the stakeholders in oncology.

The respondents included in the analysis had a mean baseline preference that is represented by an equal to moderate preference for the reference biological. No considerable difference in mean baseline preference is observed between the group that had ever prescribed, dispensed or administered a biological product and the group that had never carried out any of these actions. One respondent mentioned that he or she, or

a close family member, had experienced a serious or a life-threatening adverse reaction from a biological product. The baseline preference of this respondent for the biosimilar is lower when compared to the mean baseline preference of the whole study population. This experience could possibly explain the skeptic attitude towards biosimilars.

The findings of the current survey show differences and similarities when compared with the findings of the pilot study. It is found that in both studies a total of 7 respondents are excluded from the analysis due to a very strong or extreme baseline preference for the biosimilar. One could suggest that this are the same people as included in the pilot study. Chances are that the same people are approached because of the fact that the snowball sampling method is used. It is, however, unlikely that these 7 people are the exact same people as approached in the pilot study. First of all, the pilot study included respondents from the United States. This study only included respondents from the Netherlands. Besides, the answers obtained from the demographic questions differ in such a manner that it is improbable that the Dutch respondents included in this survey are similar to the respondents from the pilot study.

Besides this similarity between both studies, also differences were found. An explanation for the differences could be that in the pilot study also respondents from the United States were included. The experience with biosimilars between these two countries is worth mentioning, and might have caused the differences in priorities and baseline preferences. The mean baseline preference regarding biosimilars of the Dutch respondents was, however, higher when compared to the mean baseline preference of the current study. The different study sampling techniques might be a reason for these differences. The swing waiting questions were, for instance, removed from the questionnaire, since respondents indicated having difficulties answering these questions in the pilot study. Second, the pilot study used a clinical vignette in which a reference biological drugs was used of which the biosimilar already was approved. This survey, however, used reference biological products of which the biosimilar are not approved yet. The fact that the respondents don't have any previous experience with the biosimilar they had to take into account, might have caused the differences in baseline preferences. Another reason could be that the judgments regarding biosimilars in oncology have changed over time, due to, for instance, more knowledge about and experience with biosimilars. Furthermore, the types of questions asked in the pilot study differ from the types of questions asked in this survey. The pilot study used, among others, the swing weighting technique, where this study made use of a combination of the AHP and the DCE technique.

To determine the preferences of stakeholders in oncology, the multi criteria decision analysis is used. The findings of this study can be supportive in the decision making regarding the uptake of biosimilars in oncology. Using MCDA made the preferences and the underlying problems more transparent. Of the various MCDA methods, the AHP is chosen to elicit the preferences. The AHP method can be seen as a value measurement model in which direct rating is asked from the respondent. When one reflects the AHP method used in this survey, however, some limitations can be drafted. Since the first part of the questionnaire is a combination of both the AHP and the discrete choice experiments, additional calculations were required. As a result of these calculations, chances are that some detailed information has been lost regarding the preferences of the respondents. Second, using pairwise comparisons with the 1-9 reciprocal, ordinal scale could have increased the complexity of making a distinction between the answer options. Respondents could have problems with answering the questions if the criteria is, for instance, 2 or 3 times more important than the other criteria. None of the respondents indicated, however, that they had difficulties with answering the questions of the survey. Another disadvantage of the AHP is that many decision criteria and alternatives result in a large number of pairwise comparisons. This

limitation does, however, not apply for this survey since a small number of decision criteria are asked. A requirement for the decision criteria used in MCDA and AHP, is that the criteria must be judgmentally independent. The decision criteria used, i.e. effectiveness, safety and immunogenicity, are all related to each other. Even though definitions are provided of these three criteria in which the distinction is explained, chances are that the respondents didn't make a clear distinction in judging these criteria.

In this survey, the Analytic Hierarchy Process proved to be valuable in determining preferences of a variety of stakeholders under complex circumstances. The preferences of the stakeholders were clear with reference to the AHP method, and significant differences in decision criteria are found. By means of this study more insight is gained regarding the uptake of biosimilars in the Dutch oncology setting. The main limitation of this study is, however, that it has a small sample size. As a consequence, it is hard to generalize the obtained findings. Consequently, the sample size of each of the stakeholder groups is small, and the group encompassed by the term 'other' represents preferences of eight individual backgrounds. Future research is recommended to obtain a sample size that is able to generalize the findings. In addition, the fact that other criteria are mentioned by the respondents implies that here are other criteria that are important in the decision-making regarding biosimilars. Second, the respondent with a negative experience with biological products had a baseline preference which was considerably more in favor of the reference biological. More qualitative studies regarding stakeholders' attitudes and perceptions about biosimilars beyond safety, effectiveness, immunogenicity and costs savings are required. Finally, it was indicated by the respondents that they have less resistance against biosimilars when they have to threat a new patient. When they, however, have to switch the treatment of a patient who is already (effectively) treated with a biological product to biosimilar, they are considerably less positive towards biosimilars. Thus, further research should focus on switching to biosimilars during the cancer treatment.

References

1. Cancer Research UK. Worldwide cancer incidence statistics. Available from: http://www.cancerresearchuk.org/cancer-info/cancerstats/world/incidence/ Accessed 2015 March 10.

2. Ferlay J, Parkin DM, Steliarova-Foucher E. Estimates of cancer incidence and mortality in Europe in 008. Eur J Cancer 2010, 46(4):765-81.

3. WHO Media Centre. Cancer – Fact sheet. Available from: http://www.who.int/mediacentre-/factsheets/fs297/en/ Accessed 2015 March 10.

4. International Agency for Research on Cancer. Globocan 2012: Estimated cancer incidence, mortality and prevalence worldwide 2012. WHO 2012. Available from: http://globocan.iarc.fr/Default.aspx Accessed 2015 March 10.

5. American Cancer Society, Inc. The global economic costs of cancer. Available from: http://www.cancer.org/acs/groups/content/@internationalaffairs/documents/document/acspc-026203.pdf Accessed 2015 March 10.

6. National Cancer Institute (NCI). Cancer Trends Progress Report – 2011/2012 update. U.S. National Institutes of Health, 2012. Available from: http://progressreport.cancer.gov/sites/default/files/archive/report2011.pdf Accessed 2015 March 10.

7. Mariotto AB, Robin Yabroff K, Shao Y, Feuer EJ, Brown ML. Projections of the Cost of Cancer Care in the United states: 2010-2020. JNCI Journal of the National Cancer Institute 2011, 103(2): 117-128.

8. National Cancer Institute (NCI). Biological therapies for cancer. National Institutes of Health, 2013. Available from: http://www.cancer.gov/cancertopics/factsheet/Therapy/biological Accessed 2015 March 11.

9. ACS CAN. Understanding Biologic Medicines From the Cancer Patient Perspective. Available from: http://action.acscan.org/site/DocServer/ACSCAN-Biosimilars-Primer.pdf?docID=22449 Accessed 2015 March 9.

10. University of California. A short history of drug discovery. Available from: http://pharmsci.uci.edu/history.php Accessed 2015 March 27.

11. Walsh G. Biopharmaceutical benchmarks 2010. Nat Biotechnol 2010, 28(9):917-924.

12. Sekhon BS, Saluja V. Biosimilars: an overview. Dovepress J. 2011, 1(1):1-11.

13. Goel N, Stephens S. Certolizumab pegol. MAbs 2010, 2(2): 137-147.

14. Biotechnology Industry Organization (BIO). How do drugs and biologics differ? Available from: https://www.bio.org/articles/how-do-drugs-and-biologics-differ Accessed 2015 March 11.

15. Biotechnology Industry Organization (BIO). Why is patient safety a concern in the biosimilars debate? Available from: https://www.bio.org/articles/why-patient-safety-concern-biosimilars-debate Accessed 2015 March 11.

16. Samedan LTD. European Biopharmaceutical Review: spot the difference. Available from: http://www.samedanltd.com/magazine/12/issue/155/article/2980 Accessed 2015 March 11.

17. Zimney, ED. Understanding biologics: how they differ from drugs and why they cost more. Healthy Living, 2008. Available from: http://www.everydayhealth.com/columns/zimney-health-and-medical-news-you-can-use/understanding-biologics-how-they-differ-from-drugs-and-why-they-cost-more/ Accessed 2015 March 10.

18. Bennett CL, Chen B, Hermanson T, et al. Regulatory and clinical considerations for biosimilar oncology drugs. The Lancet Oncology 2014, 15(13): 594-605.

19. Hirsch B, Lyman G. Are Biosimilars Ready for Primetime in the United States? J Natl Compr Canc Netw 2011, 9: 934-942.

20. Henry D, Taylor C. Pharmacoeconomics of Cancer Therapies: Considerations with the Introduction of Biosimilars. Seminars in Oncology 2014, 41(3): 13-20.

21. Shiroiwa T, Sung YK, Fukuda T, Lang HC, Bae SC, Tsutani K. International survey on willingness-to-pay (WTP) for one additional QALY gained: what is the threshold of cost effectiveness? Health Econ 2010, 19(4): 422-37.

22. IMS Institute for Healthcare Informatics. Savings: $1 trillion over 10 years. Available from: http://www.ahipcoverage.com/wp-content/uploads/2012/08/2012-GPHA-IMS-GENERIC-SAVINGS-STUDY.pdf Accessed 2015 March 13.

23. Smith TJ, Hillner BE. Bending the Cost Curve in Cancer Care. N Engl J Med 2011, 364:2060-2065.

24. Dörner T., et al. The role of biosimilars in the treatment of rheumatic diseases. Ann Rheum Dis 2013, 72(3):322-328.

25. Bennet CL, Luminari S, Nissenson AR, et al. Pure Red-Cell Aplasia and Epoetin Therapy. N Engl J Med 2004, 351:1403-8.

26. Revers L, Furczon E. An introduction to biologics and biosimilars, part II: subsequent entry biosimilars: biosame or biodifferent? Can Pharm J 2010, 143: 184-91.

27. U.S. Food and Drug Administration (FDA). FDA Drug Safety Communication: Erythropoiesis-Stimulating Agents (ESAs): Procrit, Epogen and Aranesp. Available from: http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm200297.htm Accessed 2015 March 12.

28. Bennett CL, Cournoyer D, Carson KR, et al. Long-term outcome of individuals with pure red cell aplasia and antierythropoietin antibodies in patients treated with recombinant epoetin: a follow-up report from the Research on Adverse Drug Events and Reports (RADAR) Project. Blood 2005, 106(10): 3343-7.

29. Younes A, Gopal AK, Smith SE, et al. Results of a pivotal phase II study of brentuximab vedotin for patients with relapsed or refractory Hodgkin's lymphoma. J Clin Oncol 2012, 30(18): 2183-9.

30. Gandhi MD, Evens AM, Fenske TS, et al. Pancreatitis in patients treated with brentuximab vedotin: a previously unrecognized serious adverse event. Blood 2014, 123: 2895-2897.

31. Carson KR, Newsome SD, Kim EJ, et al. Progressive multifocal leukoencephalopathy associated with brentuximab vedotin therapy: a report of 5 cases from the Southern Network on Adverse Reactions (SONAR) project. Cancer 2014, 120(16): 2464-71.

32. European Medicines Agency (EMA). Biosimilar medical products. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Brochure/2011/03/WC500104228.pdf Accessed 2015 March 13.

33. U.S. Food and Drug Administration (FDA). Information for Consumers – Biosimilars. Available from: http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedand-Approved/ApprovalApplications/TherapeuticBiologicApplications/Biosimilars/ucm241718.htm Accessed 2015 March 13.

34. Bonig H, Becker PS, Schwebig A, Turner M. Biosimilar granulocyte-colony-stimulating factor for healthy donor stem cell mobilization: need we be afraid? Transfusion 2015, 55(2): 430-439.

35. Mulcahy AW, Predmore Z, Mattke S. The Cost Savings Potential of Biosimilar Drugs in the United States. Available from: http://www.rand.org/content/dam/rand/pubs/perspectives/-PE100/PE127/RAND_PE127.pdf Accessed 2015 March 27.

36. Ahlstrom A, King RG, Brown R, Glaudemans J, Mendelson D. Modeling Federal Cost Savings from Follow-On Biologics. Available from: http://avalere.com/research/docs/-Modeling_Budgetary_Impact_of_FOBs.pdf Accessed 2015 March 27.

37. Fox JL. Debate over details of US biosimilar pathway continues to rage. Nature Biotechnology 2012, 30: 577.

38. Grabowski H, Guha R, Salgado M. Biosimilar competition: lessons from Europe. Nature Reviews Drug Discovery 2014, 13:99-100.

39. European Medicines Agency (EMA). European Public Assessment Report (EPAR) for Aranesp. Available at: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000332-/human_med_000651.jsp&murl=menus/medicines/medicines.jsp&mid=WC0b01ac058001d125&jsenabled=true Accessed 2015 June 10.

40. European Medicines Agency (EMA). European public assessment reports. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topics/document_listing/document_listing_000318.jsp&mid=WC0b01ac0580281bf0 Accessed 2015 March 27.

41. Berghout A. Immunogenicity: impact on the design of clinical trials for biosimilars. European Generic Medicines Association 2009. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=-pages/news_and_events/events/2009/11/event_detail_000046.jsp&mid=WC0b01ac058004d5c3 Accessed 2015 March 27.

42. Bourgoin AF, Nuskey B. An outlook on US biosimilar competition. Thomson Reuters 2013. Available from: http://thomsonreuters.com/content/dam/openweb/documents/pdf/pharma-life-sciences/white-paper/anoutlookonusbiosimilarcompetition-cwp-en.pdf Accessed 2015 March 27.

43. Weise M, et al. Biosimilars: what clinicians should know. Blood 2012, 120(26): 5111-7.

44. Dranitsaris G, Amir E, Dorward K. Biosimilars of biological therapies: regulatory, clinical and commercial considerations. Drugs 2011, 71: 1527-36.

45. Blackstone EA, Fuhr JP Jr. Innovation and competition: will biosimilars succeed? Biotechnol Healthc 2012, 9: 24-7.

46. Minghetti P, et al. The regulatory framework of biosimilars in the European Union. Drug Discov Today 2012, 17(1): 63-70.

47. Schneider CK. Biosimilars in rheumatology: the wind of change. Annals of the Rheumatic Diseases 2013, 72(3): 315-318.

48. Ohly DC, Patel SK. There is no Orange Book: the coming wave of biological therapeutics. Journal of Intellectual Property Law & Practice 2011, 6(7): 464-481.

49. Kapiriri L, Norheim OF. Criteria for priority setting in health interventions in Uganda. Exploration of stakeholders' values. Bull World Health Organ 2004, 82: 172-9.

50. Mills A, Bennett S, Bloom G, Gonzalez-Block MA, Pathmanathan I. Strengthening health systems: the role and promise of policy and systems research. Alliance for Health Policy and Systems Research/Alliance for Health Policy and Systems Research (AHPSR) 2004. Available from: http://www.who.int/alliance-hpsr/resources/Strengthening_complet.pdf Accessed 2015 April 1.

51. Guindo LA, Wagner M, Baltussen R, Rindress D, van Til J, Kind P, Goetghebeur MM. From efficacy to equity: literature review of decision criteria for resource allocation and healthcare decision making. Cost Eff Resour Alloc 2012, 10(1):9.

52. Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. Cost Eff Resour Alloc 2006, 4:14.

53. Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D: Bridging health technology assessment (HTA) and efficient health care decision making with multicriteria decision analysis (MCDA): applying the EVIDEM framework to medicines appraisal. Med Decis Mak 2012, 32(2):376-388.

54. Hummel M, IJzerman MJ. The past and future of the AHP in health care decision making. Available from: http://doc.utwente.nl/79775/1/past_and_future.pdf Accessed 2015 April 1.

55. Williams P, Mauskopf J, Lebiecki J, Kilburg A. Using multicriteria decision analysis during drug development to predict reimbursement decisions. Journal of Market Access & Health Policy 2014, 2:25270 http://dx.doi.org/10.3402/jmahp.v2.25270

56. Walker S, Phillips L, Cone M. Benefit risk assessment model for medicines: Developing a structured approach to decision making. Epsom: Center for Medicines Research International, Institute for Regulatory Science 2006.

57. Mussen F, Salek S, Walker S. A quantitative approach to benefit-risk assessment of medicines – Part 1: The development of a new model using multi-criteria decision analysis. Pharmacoepidemiol Drug Saf. 2007, 16:S2-S15.

58. Thokala, P., & Duenas, A. Multiple Criteria Decision Analysis for Health Technology Assessment. Value in Health 15 2012. http://dx.doi.org/10.1016/j.jval.2012.06.015

59. Burgha, C.M. Structure of Multi-Criteria Decision-Making. The Journal of the Operational Research Society 2004, 55(11): 1156-1168.

60. Liberatore MJ, Nydick RL. The analytic hierarchy process in medical and health care decision making: a literature review. Eur J Oper Res 2008, 189: 194-207.

61. van Til, J., et al. Does technique matter; a pilot study exploring weighting techniques for a multi-criteria decision support framework. Cost Effectiveness and Resource Allocation 2014, 12:22

62. Reynolds T, Jolly J. Measuring personal values: an evaluation of alternative methods. J Marketing Sci 1980, 17:531-6.

63. Saaty, TL, Vargas LG. The possibility of group choice: pairwise comparisons and merging functions. Soc Choice Welf 2012, 38: 481-496.

64. Figuaera J, Greco S, Ehrgott M. (Eds) (2005). Multiple Criteria Decision Analysis, State of the Art Surveys, New York: Springer.

65. Thanesuen S, Kagaya S, Uchida K. Application of Analytic Hierarchy Process on Preferable Speed Limit for Logistics Company: A Case Study on Hokkaido Roads. Canadian Journal of Transportation**,** [S.l.], v. 1, n. 2, jul. 2007. ISSN 1715-2127. Available at: <http://journals.hil.unb.ca/index.php/CJT/article/view/654/4739>. Accessed 2015 April 2.

66. Saaty, T.L. (1988). Multicriteria Decision Making: The Analytic Hierarchy Process, McGraw-Hill, New-York, 1980. Republished by University of Pittsburgh, Pittsburgh.

67. Saaty, T.L. (1980) The Analytic Hierarchy Process, New York: McGraw Hill. International, Translated to Russian, Portuguese, and Chinese, Revised editions, Paperback (*1996, 2000*), Pittsburgh: RWS Publications.

68. Saaty, T.L. and Vargas, L.G. (2000). Models, Methods, Concepts and Applications of the Analytic Hierarchy Process, Bosten: Kluwer Academic Publishers.

69. Dolan JG. Shared decision-making – transferring research into practice: The Analytical Hierarchy Process (AHP). Patient Educ Couns. 2008, 73: 418-425.

70. Bana e Costa, Carlos A. and Vansnick, Jean-Claude (2001). A critical analysis of the eigenvalue method used to derive priorities in AHP. Operational Research working papers, LSEOR 01.42. Department of Operational Research, London School of Economics and Political Science, London, UK.

71. Danner M, et al. Integrating patients' views into health technology assessment: Analytic hierarchy process (AHP) as a method to elicit patient preferences. International Journal of Technology Assessment in Health Care 2011, 27 (4): 369-375.

72. Saaty TL. Decision making with the Analytic Hierarchy Process. Int J Serv Sci. 2008, 1:83-98.

73. Yau, Y. Multi-criteria decision making for urban built heritage conservation: application of the analytic hierarchy process.' Journal of Building Appraisal 2009, 4: 191-205.

74. Saaty TL. Decision-making with the AHP: Why is the principal eigenvector necessary? European Journal of Operational Research 2013, 145(1): 85-91.

75. Hummel JM, Steuten LG, Groothuis-Oudshoorn CJ, Mulder N, IJzerman MJ. Preferences for Colorectal Cancer Screening Techniques and Intention to Attend: a Multi-Criteria Decision Analysis. Appl Health Econ Health Policy 2013, 11(5): 499-507.

76. Azevedo VF, Felippe LR, Machado DG. Opinion of some Brazilian rheumatologists about biosimilars. Rev. Bras. Reumatol. Vol. 51 no. 6 São Paulo (2011). http://dx.doi.org/10.1590/S0482-50042011000600013

77. Biernacki P, Waldorf D. Snowball Sampling: Problems and Techniques of Chain Referral Sampling. Sociological Methods & Research 1981, 10: 141-163.

78. Manko J, et al. A clinical comparison of the efficacy and safety of biosimilar G-CSF and originator G-CSF in haematopoietic stem cell mobilization. Pharmacol Rep 2014, 66(2): 239-242.

79. Publicover A, et al. Use of a biosimilar granulocyte colony-stimulating factor for peripheral blood stem cell mobilization: an analysis of mobilization and engraftment. Br J Haematol 2013, 162(1):107-111.

80. Remenyi P, et al. Peripheral Blood Stem Cell Mobilization and Engraftment after Autologous Stem Cell Transplantation with Biosimilar rhG-CSF. Adv Ther 2014, 31(4):451-460.

81. Schmitt M, et al. Mobilization of PBSC for allogeneic transplantation by the use of the G-CSF biosimilar XM02 in healthy donors. Bone Marrow Transplant 2013, 48(7):922-925.

82. Sivgin S, et al. The comparison of Filgrastim (Neupogen®), biosimilar filgrastim (Leucostim®) and Lenograstim (Granocyte®) as a first line peripheral blood stem cell mobilization strategy in autologous hematopoieitic stem cell transplantation: a single center experience from Turkey. Transfus Apher Sci. 2013, 48(3):315-320.

83. Yafour N, Brahimi M, Osmani S, et al. Biosimilar G-CSF (filgrastim) is effective for peripheral blood stem cell mobilization and non-cryopreserved autologous transplantation. Transfus Clin Biol 2013, 20:502-4.

84. Gascón P, Tesch H, Verpoort K, et al. Clinical experience with Zarzio® in Europe: what have we learned? Support Care Cancer 2013, 21:2925-32.

85. Ianotto J, Ngo Sack F, Couturier M, Tempescul A, Mugnier N, Delepine P, et al. Biosimilars of filgrastim in autologous stem cell transplant: reduction in granulocyte-colony stimulating factor costs, but similar effects on bone marrow recovery. Leuk Lymphoma 2014, 55: 74–77.

86. Kerkhofs L, Boschetti G, Lugini A, Stanculeanu DL, Palomo AG. Use of biosimilar epoetin to increase hemoglobin levels in patients with chemotherapy-induced anemia: real-life clinical experience. Future Oncol. 2012, 8(6):751-756.

87. Sagara Y, et al. The Efficacy and Safety of FSK0808, Filgrastim Biosimilar: A Multicenter, Non-randomized Study in Japanese Patients with Breast Cancer. Jpn J Clin Oncol. 2013, 43(9):865-873.

88. Salesi N, et al. Biosimilar medicines in oncology: single-center experience with biosimilar G-CSF. Future Oncol. 2012, 8(5):625-30.

89. Schmitt M, Publicover A, Orchard K, Görlach M, Wang L, Schmitt A, et al. Biosimilar G-CSF based mobilization of peripheral blood hematopoietic stem cells for autologous and allogeneic stem cell transplantation. Theranostics 2014, 4: 280–289.

90. Verpoort K, Mohler TM. A non-interventional study of biosimilar granulocyte colony-stimulating factor as prophylaxis for chemotherapy-induced neutropenia in a community oncology center. Ther Adv Med Oncol. 2012, 4(6):289-293.

91. Castelli R, et al. Biosimilar epoetin in elderly patients with low-risk myelodysplastic syndromes improves anemia, quality of life, and brain function. Ann Hematol. 2014, 93(9):1523-9.

92. FirstWord Pharma. FirstWord Lists – Pharma's 50 biggest selling drugs: AbbVie's Humira joins the $10 billion club. Available from: http://www.firstwordpharma.com/node/1194000#axzz3ZTMG2xyD Accessed 2015 June 29.

93. BPMSG. New AHP Excel template with multiple inputs. Available from: http://bpmsg.com/new-ahp-excel-template-with-multiple-inputs/ Accessed 2015 June 29.

94. ASCO. Cost of Cancer Drugs Should be Part of Treatment Decisions. Available from: https://am.asco.org/cost-cancer-drugs-should-be-part-treatment-decisions Accessed 2015 June 26.

95. European Medicines Agency (EMA). Summary of product characteristics. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000406/WC500022207.pdf Accessed 2015 June 30.

96. U.S. Food and Drug Administration (FDA). Gleevec (Imantinib Mesylate) Questions and Answers. Available from: http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationfor-PatientsandProviders/ucm110505.htm Accessed 2015 June.

97. Cancer Research UK. Bevacizumab (Avastin). Available from: http://www.cancerresearchuk.org/about-cancer/cancers-in-general/treatment/cancer-drugs/bevacizumab Accessed 2015 June 30.

98. Semenza, GL. A new weapon for attacking tumor blood vessels. N. Engl. J. Med. 2008, 358(19):2066-7.

99. Sliesoriaitis S, Twafik B. Bevacizumab-induced bowel perforation. The journal of the American OSteopatic Associaton 2011, 111(7):437-44.

100. Schmucker, C., et al. A safety review and meta-analyses of Bevacizumab and Ranibizumab: Off-Label versus Goldstandard. PLoS ONE 2012, 7(8): e42701.

101. Niederwieser D, Schmitz S. Biosimilar agents in oncology/haematology: from approval to practice. European Journal of Haematology 2011, 86(4): 277-288.

102. Bren L. Cancer drugs: weighing the risks and benefits. Available from: http://www.cancertreatmentwatch.org/general/drugs.shtml Accessed 2015 June 30.

103. Harvard T.H. Chan. Prevalence and Incidence Defined. Available from: http://www.hsph.harvard.edu/obesity-prevention-source/prevalence-incidence/ Accessed 2015 March 10.

104. Oxford University Press. Mortality. Available from: http://www.oxforddictionaries.com/definition/english/mortality Accessed 2015 March 10.

Appendix A: Questionnaire

Survey Regarding the Use of Biosimilars Compared to Use of Reference Biological Drugs

Dennis W. Raisch, PhD[1,2], Maarten Ijzermann[2], PhD; Charles L. Bennett, MD, PhD[3]

[1] University of New Mexico, College of Pharmacy

[2] Department of Health Technology and Services Research, University of Twente

[3] The Southern Network on Adverse Reaction (SONAR) project, the South Carolina Center of Economic Excellence for Medication Safety, University of South Carolina College of Pharmacy

This survey is to gather your perceptions about reference biologicals compared to biosimilars, which can be marketed after a biological product reaches its patent expiration date.

This survey consists of three parts: (1) questions about your preferences with regard to reference biologicals and biosimilars, (2) questions about different criteria in the decision-making process that are important to you, and (3) questions about your demographics.

We are asking about your opinions. There are no right or wrong answers. All your responses are anonymous.
It takes about 15 minutes to complete the questionnaire. Thank you very much for your time and effort!

Before the questionnaire starts, a few facts about biosimilars will be presented:

- At $170 billion in sales in 2012, biologicals are the fastest growing pharmaceutical products and account for 5 of the top 10 products in terms of sales.
- Once a **reference biological product** (the originator) reaches its patent expiration date, guidelines for approval of **biosimilars** have been established by the Food and Drug Administration, European Medicines Agency, and World Health Organization. The guidelines require at least one equivalence (or non-inferiority) study compared to the reference biological, with or without a placebo arm.
- **Biosimilars** are uniquely bioengineered and are not exact copies of the **reference biological**, but are considered equivalent in terms of efficacy, safety, immunogenicity, pharmacokinetics, and pharmacodynamics.
- Equivalence in **safety** means that the **biosimilar** is similar to the **reference biological** in frequency and severity of adverse events.
- Equivalence in **immunogenicity** means that the **biosimilar** is similar to the **reference biological** in frequency and severity of immune response.
- The marketing approval for a **biosimilar** is granted based upon studies typically performed in a smaller number of patients compared to the number required for approval of the **reference biological** application.
- Post-marketing effectiveness and safety, including immunogenicity, studies for **biosimilars** may be required by regulatory agencies on a case-by-case basis. Post-marketing studies have larger sample sizes and provide additional evidence of equivalence through observational methods (such as patient registries) and/or clinical trials.
- **Biosimilars** are priced less than the **reference biological**, so their use may reduce treatment costs.

For the first 7 questions, *please consider this scale:*

Value	Meaning
1	Equally preferred

3	Moderately preferred
5	Strongly preferred
7	Very strongly preferred
9	Extremely preferred
2,4,6,8	Intermediate values

The first part includes 8 questions about comparing your preferences of reference biologicals and biosimilars. Each of these 8 questions will be provided with a graphic. The graphic looks like this:

Approval Data	Post-marketing Effectiveness Data	Post-marketing Safety Data	Post-marketing Immunogenicity Data	Cost is 50% less

This graphic depicts which information is available for the **biosimilar** with green shading. This green shading shows the available information for the biosimilar in that particular question. Assume that the following graphic is provided:

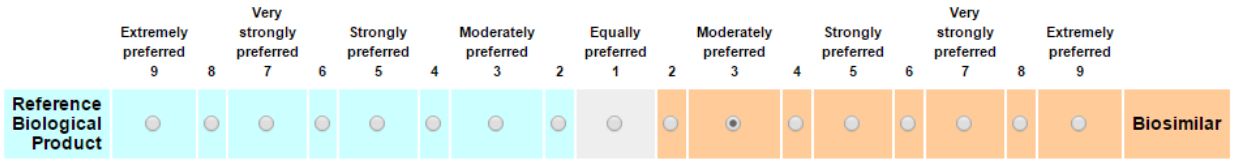
Approval Data	Post-marketing Effectiveness Data	Post-marketing Safety Data	Post-marketing Immunogenicity Data	Cost is 50% less

This means that the approval data and the post-marketing effectiveness data is available for the biosimilar in this question. Assume that the data without the green shading is **not** available or applicable to the question.

For the next group of questions an example will be provided. The questionnaire will start after this example.

Assume in this survey that the **approval data** shows evidence in a limited number of patients, that the biosimilar performs **equally well** in terms of effectiveness, safety and immunogenicity as the reference biological.

The item below is an example where someone has indicated that, given approval data only for the biosimilar, he/she moderately prefers to use a biosimilar over a reference biological:



1. General question: Given a patient who may benefit from a reference biological and there is a newly-marketed ***biosimilar available at 25% lower cost***. The ***approval data*** shows that the biosimilar performs equally well as the reference biological in terms of effectiveness, safety and immunogenicity. Would you prefer using the reference biological or the biosimilar? And to what extent do you prefer this decision?

Approval Data Only	Post-marketing Effectiveness Data	Post-marketing Safety Data	Post-marketing Immunogenicity Data	Cost is 50% less

Reference	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Biosimilar

Biological	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o

2. Which of these reference biological products do you most often prescribe for cancer treatment/are you the most familiar with?

    - o CD20 antigen inhibitor, e.g. Rituximab (Rituxan, Mabthera)
    - o Human epidermal growth factor receptor 2 (HER-2) inhibitor, e.g. Trastuzumib (Herceptin)
    - o Human vascular endothelial growth factor inhibitor, e.g. Bevacizumab (Avastin)
    - o Tyrosine kinase inhibitor, e.g. Imantinib (Gleevac)
    - o Epidermal growth factor receptor inhibitor, e.g. Cetuximab (Erbitux)

You have selected: ____.

In answering the upcoming questions, consider a "typical" oncology patient who you treat with the biological product you selected. The patient has no contraindications, physiological abnormalities, or co-morbidities that might preclude its use. Assume that a biosimilar for the reference biological has been approved by the FDA and the EMA and is available (at lower cost) in your practice setting.

Unless specified in the question, assume that only approval data is available. In particular: there is **no** post-marketing information on **effectiveness, safety or immunogenicity**.

3. Assuming that only approval data exist and ***the biosimilar cost is 25% less*** than the reference biological, would you prefer using the reference biological or the newly-marketed biosimilar*?* And to what extent do you prefer using the reference biological or the biosimilar?

Approval Data Only	Post-marketing Effectiveness Data	Post-marketing Safety Data	Post-marketing Immunogenicity Data	Cost is 50% less

Reference Biological	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Biosimilar

4. Assuming that only approval data exist and the ***biosimilar costs 50% less*** than the reference biological would you prefer using the reference biological or the newly-marketed biosimilar*?* And to what extent do you prefer using the reference biological or the biosimilar?

Approval Data Only	Post-marketing Effectiveness Data	Post-marketing Safety Data	Post-marketing Immunogenicity Data	Cost is 50% less

Reference Biological	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Biosimilar

5. Assuming that approval data exist plus a ***post-marketing study has shown that the biosimilar has equal effectiveness (and cost is 25% less)***, compared to the reference biological would you prefer using the reference biological or the biosimilar*?* And to what extent do you prefer using the reference biological or the biosimilar?

Approval Data	Post-marketing Effectiveness Data	Post-marketing Safety Data	Post-marketing Immunogenicity Data	Cost is 50% less

Reference Biological	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Biosimilar

6. Assuming that approval data exist, plus a ***post-marketing study has shown that the biosimilar has equal safety (and cost is 25% less)*** compared to the reference biological would you prefer using the reference biological or the biosimilar*?* And to what extent do you prefer using the reference biological or the biosimilar?

Approval Data	Post-marketing Effectiveness Data	Post-marketing Safety Data	Post-marketing Immunogenicity Data	Cost is 50% less

Reference Biological	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Biosimilar

7. Assuming that approval data exist, plus a ***post-marketing study has found that the biosimilar has equal immunogenicity (and cost is 25% less)*** compared to the reference biological; would you prefer using a reference biological or the biosimilar*?* And to what extent do you prefer using the reference biological or the biosimilar?

Approval Data	Post-marketing Effectiveness Data	Post-marketing Safety Data	Post-marketing Immunogenicity Data	Cost is 50% less

Reference Biological	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Biosimilar

8. Assuming the ***biosimilar costs 25% less*** than the reference and is ***equal in efficacy, safety, and immunogenicity based upon post-marketing research,*** would you prefer using the reference biological or the biosimilar*?* And to what extent do you prefer using the reference biological or the biosimilar?

Approval Data	Post-marketing Effectiveness Data	Post-marketing Safety Data	Post-marketing Immunogenicity Data	Cost is 50% less

Reference Biological	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Biosimilar

9. Assuming the ***biosimilar costs 50% less*** than the reference and is ***equal in efficacy, safety, and immunogenicity based upon post-marketing research,*** would you prefer using the reference biological or the biosimilar? And to what extent do you prefer using the reference biological or the biosimilar?

Approval Data	Post-marketing Effectiveness Data	Post-marketing Safety Data	Post-marketing Immunogenicity Data	Cost is 50% less

Reference Biological	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Biosimilar

You have finished part 1. For the questions of part 2, you are asked to choose between two different criteria. There are no right or wrong answers, we are asking about factors that are important to you personally. Good luck with filling in the questions!

Criteria weighting questions

Assessing relative importance of criteria

10. What **post marketing data** would be more important to you: evidence on the **effectiveness** or evidence on the **safety** of the biosimilar? And to what extent do you consider this criteria to be more important?

Effectiveness	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Safety

11. What **post marketing data** would be more important to you: evidence on the **effectiveness** or evidence on the **immunogenicity** of the biosimilar? And to what extent do you consider this criteria to be more important?

Effectiveness	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Immunogenicity

12. What **post marketing data** would be more important to you: evidence on the **safety** or evidence on the **immunogenicity** of the biosimilar? And to what extent do you consider this criteria to be more important?

Safety	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Immunogenicity

13. What would be more important to you: **post marketing data on effectiveness** or **cost savings** from the biosimilar? And to what extent do you consider this criteria to be more important?

Post-marketing Effectiveness	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Cost Savings

14. What would be more important to you: **post marketing data on safety** or **cost savings** from the biosimilar? And to what extent do you consider this criteria to be more important?

Post-marketing Safety	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Cost Savings

15. What would be more important to you: **post marketing data on immunogenicity** or **cost savings** from the biosimilar? And to what extent do you consider this criteria to be more important?

Post-marketing Immunogenicity	9 o	8 o	7 o	6 o	5 o	4 o	3 o	2 o	1 o	2 o	3 o	4 o	5 o	6 o	7 o	8 o	9 o	Cost Savings

16. If applicable, please list any other important criteria in deciding whether to use a biosimilar instead of a reference biological product:
    _____
    _____
    _____

You have also finished part 2. In the last part of the survey, only a few simple questions will be asked about your demographics. After these questions, the survey is finished. Good luck with the final questions!

Demographics and Background

17. Please check the range that includes your age (in years):

    - o <30 years
    - o 31-40
    - o 41-50
    - o 51-60
    - o >60 years
    - o No answer

18. Please check your gender:

    - o Male
    - o Female

19. Please check the role that best describes your primary activity:

    - o Physician

- Oncologist
- Pharmacist
- Nurse
- Other: _____
- No answer

20. What is your primary activity?

    - Clinical care provider
    - Involved in health care formulary decisions
    - Involved in benefits structure for health payer
    - Government policy maker
    - Researcher (clinical trial, basic science, health services)
    - Pharmaceutical company (any role)
    - Teaching at a University or Medical Facility
    - Other: _____
    - No answer

21. Have you ever prescribed, dispensed, or administered a biological product?

    - Yes
    - No
    - No answer

22. Have you ever received a biological product as a patient?

    - Yes
    - No
    - No answer

23. Have you or has a close family member ever experienced a serious or life threatening adverse reaction from a drug or biological product?

    - Yes
    - No
    - No answer

24. Please provide any comments or suggestions you have about the survey:

_____
_____
_____

The questionnaire is finished. Thank you very much for your cooperation!

If you know individuals who could complete this questionnaire based on their background, I kindly ask you to forward the following link to them:

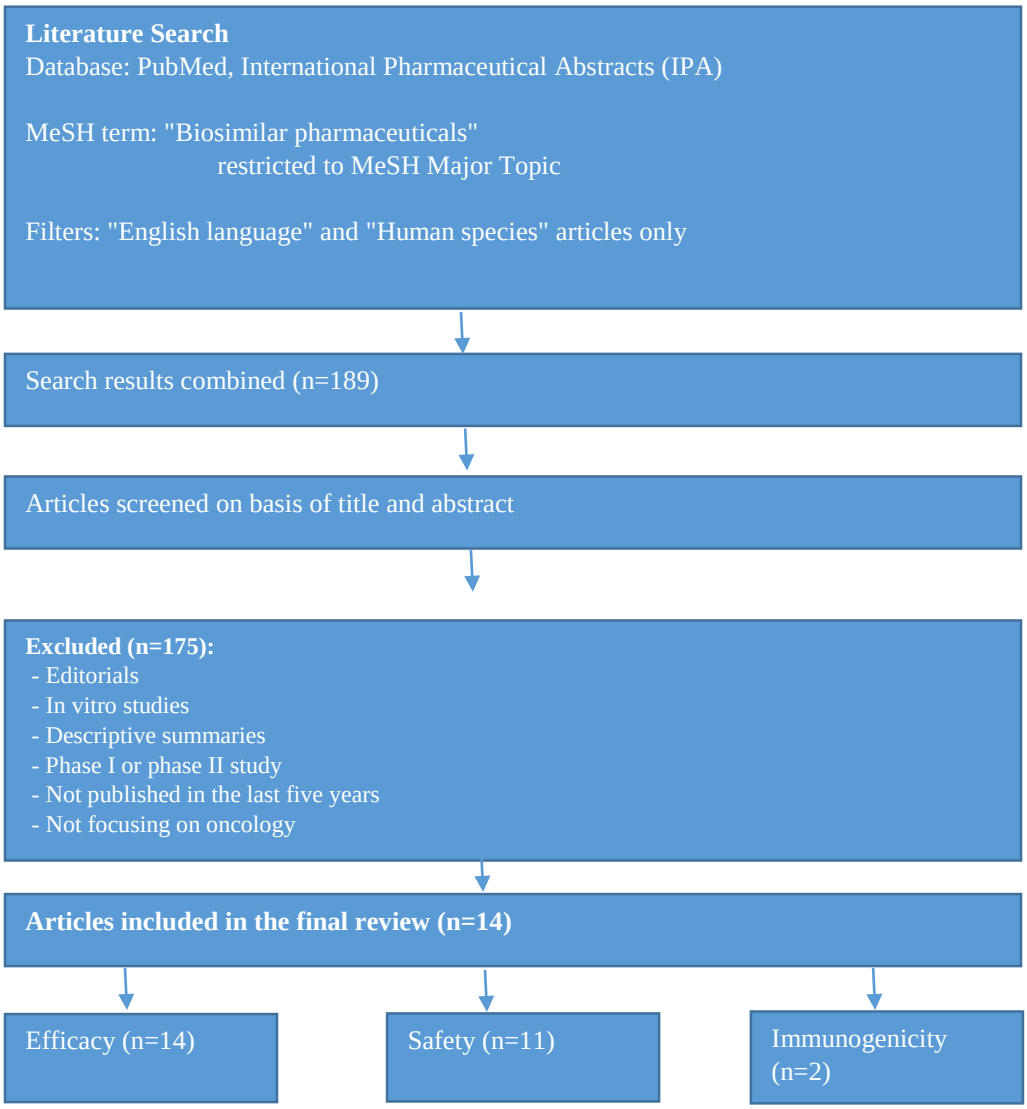
https://surveys-igs.utwente.nl/index.php?r=survey/index/sid/739124/lang/en

Appendix B: Literature search

The systematic literature review is carried out by performing a literature search in the databases PubMed and International Pharmaceutical Abstracts (IPA). The following search term is used in both databases to find articles:

("Biosimilar Pharmaceuticals"[Majr] AND ("humans"[MeSH Terms] AND English[lang]))

The articles found were screened on the basis of the title and the abstract, to identify articles that were eligible for the review. Criteria to include articles were: (1) the study design must provide a comparison of biosimilars, (2) the result section must include data, (3) the article must focus upon oncology and (4) the study must be published in the last five years. If articles were excluded, the reasons were: (1) the article was an editorial, (2) it was an in vitro study, (3) it was a descriptive summary, (4) it was a phase I or phase II study, or (5) the article could not be located.



Appendix C: Results of the individual studies

Manko, J., et al. *Pharmacol Rep.* 2014;66(2): 239-242.

Drug	G-CSF (Zarzio®)
Study design	Randomized trial
Number of participants	108
Treatment	Patients (Multiple myeloma, non-Hodgkin's or Hodgkin's lymphoma)
Effectiveness, safety and/or immunogenicity	Effectiveness, safety
Main results	Originator group: 95% didn't mobilize enough CD34+ cells, Biosimilar group: 11% didn't mobilize enough CD34+ cells, Adverse events were comparable between both groups, Safety and efficacy were comparable between both groups.
Negative findings	No time-frame was mentioned
Positive findings	A control group was used, Wide range (age) was used (19-69 years), Multiple diagnoses were taken into account, Both female and male patients included.

Publicover, A., et al. *Br J Haematol.* 2013;162(1):107-111.

Drug	G-CSF (Ratiograstrim ®)
Study design	Observational
Number of participants	285
Treatment	Patients (Peripheral blood stem cell harvest)
Effectiveness, safety and/or immunogenicity	Effectiveness
Main results	Intervention group: N=154, Control group: N=131, Differences between both groups were not significant, The use of G-CSF can result in cost savings.
Negative findings	In both the originator and the biosimilar group some harvest were 'inadequate' (respectively 12 and 16%).
Positive findings	A control group was used (received Neuprogen®), A large time period (data of 3 years included), The control group received the originator.

Remenyi, P., et al. Adv Ther. 2014;31(4):451-460.

Drug	G-CSF (rhG-CSF)
Study design	Observational
Number of participants	110
Treatment	Patients (Hematological/lymphoid malignancies)
Effectiveness, safety and/or immunogenicity	Effectiveness, safety
Main results	Study I: N=70; Study II: N=40, The median duration to (neutrophil/leukocyte/platelet) engraftment was comparable between both studies, The biosimilar and the originator can be compared in terms of both kinetics of PBSC mobilization and the proceeds of CD34+, The use of the biosimilar can be seen as safe and effective.
Negative findings	Study I consisted of more male patients (60%), Study II consisted of more female patients (57,5%).
Positive findings	Patients from two medical centers included, Two studies compared to each other.

Schmitt, M., et al. *Bone Marrow Transplant.* 2013;48(7):922-925.

Drug	G-CSF (XM02)
Study design	Observational
Number of participants	44
Treatment	Patients/donors (PBSC mobilization)
Effectiveness, safety and/or immunogenicity	Effectiveness, safety
Main results	Group 1: 22 patients, Group 2: 22 donors, In both groups: 54.5% donors (N=6) reported arthralgia's (expected), Effectiveness and safety aspects are comparable between the biosimilar and the originator group, Incidence of graft rejection and possible side effects/adverse events was comparable between both groups.
Negative findings	Small sample size (N=44), Only investigated short-term safety effects
Positive findings	Two groups are compared to each other, Engrafting was 100%.

Sivgin, S., et al. *Transfus Apher Sci*. 2013;48(3):315-320.

Drug	G-CSF (Neupogen, Leucostim)
Study design	Observational
Number of participants	96
Treatment	Patients (autoHSCT)
Effectiveness, safety and/or immunogenicity	Effectiveness
Main results	The amount of PB CD34+ cells was higher when patients received Leucostim compared with the Granocyte group, Leucostrim can be compared with Neupogen in terms of PBSC mobilization.
Negative findings	Single center experience, Much more male patients included (71,8%).
Positive findings	Multiple groups compared with each other, Differences between groups of agents were not significant (p=0.067).

Yafour, N., et al, *Transfus Clin Biol*. 2013;20(5-6);502-4.

Drug	G-CSF (Zarzio®)
Study design	Observational
Number of participants	20
Treatment	Patients (Hematological malignancies)
Effectiveness, safety and/or immunogenicity	Effectiveness, safety
Main results	Group 1 (Biosimilar): N=10 / Group 2 (Neupogen): N=10, The groups that are compared were comparable with regard to characteristics, apheresis, engraftment and possible side effects/adverse events.
Negative findings	Sample size is small, 80% of the participants was male, No data available from RCTs.
Positive findings	Two groups were compared, Findings are consistent with other, comparable research regarding biosimilar Zarzio®.

Iannoto, J.C., et al. *Leuk Lymphoma.* 2014;55(1):74-77.

Drug	G-CSF
Study design	Observational
Number of participants	115
Treatment	Patients (Lymphoma/myeloma, receiving ASCT)
Effectiveness, safety and/or immunogenicity	Effectiveness, safety
Main results	Group 1 (intervention): N=65 / Group 2 (control): N = 50, No differences found between the biosimilar and the originator in terms of safety and effectivity, The parameters observed were "less favorable in patients with lymphoma than in the patients with myeloma", The use of biosimilar G-CSF can result in significant cost reductions, No reduction in hospitalization costs were noticed.
Negative findings	More cases of bacteremia found in the lymphoma group that received the biosimilar (although not significant), A randomized study is needed to verify the obtained results.
Positive findings	Two groups were compared, The characteristics between the two groups were comparable, The effects of the use of biosimilar G-CSF were observed using 19 parameters

Gascon, P., et al. *Support Care Cancer.* 2013;21:2925-32.

Drug	G-CSF (Zarzio®)
Study design	Observational
Number of participants	1302 (49% breast; 17% lung; 15% blood cancer).
Treatment	Patients (oncology patients with neutropenia after receiving cytotoxic therapy)
Effectiveness, safety and/or immunogenicity	Effectiveness, safety, immunogenicity
Main results	29 patients (2,2%) got an episode of febrile neutropenia and 104 patients (8,5%) experienced grade four neutropenia = is in range with previous studies, Zarzio ® is effective in preventing neutropenia that is induced by chemotherapy, No remarkable safety issues occurred while using Zarzio ®, Similar safety profile (8% bone pain versus 22% bone pain in the reference biological), Immunogenicity: no antibodies were found that could result in neutralizing.
Negative findings	All included studies were not interventional, Of 16% of the patient it was unknown which treatment they had received, In two studies there were some disturbances regarding chemotherapy regimens (10% and 7% respectively).
Positive findings	Pooled analysis of 5 post-marketing studies, The included studies reported the clinical use of Zarzio ® in the real-life setting, Studies from 12 different EU countries, 3 single-center studies and 2 multi-center studies, All adult patients included (18 years and older),

	Multiple different cancer types included.

Sagara, Y., et al. *Jpn J Clin Oncol.* 2013;43(9):865-873,

Drug	G-CSF (FSK0808)
Study design	Non-randomized study
Number of participants	104
Treatment	Patients (Breast cancer)
Effectiveness, safety and/or immunogenicity	Effectiveness, safety, Immunogenicity
Main results	Average time duration of neutropenia: 2.2±1.5 days, Febrile neutropenia (incidence): 34.6%, Multiple adverse drug reactions occurred in >5% of the patients, Back pain was the most common adverse drug reaction (60,6%), No anti-G-CSF antibodies found before and after treatment with the biosimilar (N=102).
Negative findings	Non-randomized setting (no control group consisting of patients receiving the originator used), A total of 20 patients excluded from the analysis (efficacy) due to various reasons, 104/104 patients (100%) experienced an AE, A total of 1795 AE reported.
Positive findings	Results are comparable with findings regarding AE of the originator.

Salesi, N., et al. *Future Oncol.* 2012:8(5):625-30.

Drug	G-CSF (Zarzio®)
Study design	Observational
Number of participants	48
Treatment	Patients (Solid tumors)
Effectiveness, safety and/or immunogenicity	Effectiveness, safety
Main results	Group 1 (primary prophylaxis): N=37 / Group 2 (secondary prophylaxis): N=11, 3 patients experienced febrile neutropenia (= 6,25%), 6 patients experienced nonfebrile (grade four) neutropenia (=12,5%), According to the findings, biosimilar G-CSF can be seen as both safe and effective in patients suffering from neutropenic complications.
Negative findings	The length of administration (median) was only seven days, Difference in the sample size per group is relatively large, Single-centered and small sample size, No control group (originator) was used.
Positive findings	Two groups compared to each other, Febrile neutropenia can be effectively treated with antibiotics, no hospitalization necessary.

Verpoort, K. and Mohler, T.M. *Ther Adv Med Oncol.* 2012;4(6):289-293.

Drug	G-CSF (Zarzio ®, Filgrastim Hexal ®)
Study design	Observational
Number of participants	102
Treatment	Patients (various types of cancer)
Effectiveness, safety and/or immunogenicity	Effectiveness, safety
Main results	Group 1 (biosimilar): N=77 / Group 2 (originator): N=25, Biosimilar can be compared to the originator in terms of clinical outcomes, No safety problems occurred, In 91% of the patients the usage of the biosimilar resulted in preventing dose reductions and discontinuation, 6,5% of the patients needed a dose reduction, 2,5% of the patients needed to discontinue.
Negative findings	Only patients from one center (single center), 77 patients received the biosimilar, versus 25 patients who received the originator, Mostly elderly (median age is 67 years).
Positive findings	Patients from a large community practice specialized in oncology were included, The age of the included patients differed between 20-83 years, Multiple types of cancer included, 2.5-year time period, Two groups are compared with each other.

Castelli R., et al. *Ann Hematol.* 2014 Sep;93(9):1523-9.

Drug	Epoetin (Binocrit)
Study design	Observational
Number of participants	24
Treatment	Patients (MDS patients)
Effectiveness, safety and/or immunogenicity	Effectiveness
Main results	14 male patients, 10 female patients, 16 of the 24 patients got an erythroid response (67%), 15 of the 24 patients got independent from transfusion (62,5%), 7 of the 24 patients did not respond to the biosimilar (29,1%), The values of Hb were much higher (in the group of patients that responded to the biosimilar) after treatment compared with the Hb level before treatment ($p<0.001$), No single patient experienced major adverse/side effects.
Negative findings	Only tested in patients with the age of >65 years (average 72 years), No control group was used (to compare with the originator ESA), Short-term (12 weeks of therapy, follow-up again 12 weeks), Small sample size.
Positive findings	Significant, positive relation between Hb-improvement and the variety in FACT-An scores,

	Positive correlation between ER and improved cognitive functions + QoL in patients with MDS, Only newly diagnosed MDS patients were included, Responders rate is comparable with already conducted, comparable data regarding the originator drug.

Kerkhofs, L., et al. *Future Oncol.* 2012;8(6): 751-756.

Drug	Epoetin (Binocrit)
Study design	Observational
Number of participants	113
Treatment	Patients (cancer, and anemia as a result of chemotherapy)
Effectiveness, safety and/or immunogenicity	Effectiveness, safety, immunogenicity
Main results	79% of the patients had a Hb response, Response rates comparable between patients that received a dose of 30,000 in comparison with the group that received a dose of 40,000 IU/week, Hb response rate increased when patients received iv. Iron in comparison with patients that did not received iv. Iron (93% vs. 77%; significant), No serious adverse events reported.
Negative findings	No control group was used
Positive findings	Five European centers included (Netherlands, France, Italy, Romania, and Spain).

Schmitt, M., et al. *Theranostics*. 2014;4:280-9.

Drug	G-CSF
Study design	Comprehensive review
Number of participants	904
Treatment	Patients (Hematological malignancies)
Effectiveness, safety and/or immunogenicity	Effectiveness, safety, immunogenicity
Main results	No differences according to post transplantation engraftments between originator/biosimilar group (effectiveness is comparable), No differences regarding to side effects between originator/biosimilar group, No increase in toxicity found, Immunogenicity: differences between originator/biosimilar groups were not significant.
Negative findings	No summarization of the data/data analysis included, Chances are that there is a conflicts of interest among the authors, Non-randomized setting (no control group used).
Positive findings	Sample size is relatively large (N=904), If possible, siblings were included in the research.

Appendix D: Literature review effectiveness outcome measures

WBC Count/leukocytes/hemoglobin/platelets ($x10^9$/L):

Author, year	(Primary) Outcome Measure(s)	Outcome Biosimilar	Outcome Reference Biological	P-value
Iannotto et al., 2014[85]	*BM recovery (Lymphoma):*			
	Leukocytes ($x10^9$/L)	4.6 (min-max: 1-7.6),	4 (min-max: 1.9-8.2)	0.17
	Hemoglobin (g/dL)	11.5 (min-max: 9.1-15.7)	11.3 (min-max: 9.1-13.3)	0.44
	Platelets ($x10^9$/L)	195 (min-max: 75-429)	207.9 (min-max: 45-426)	0.53
	BM recovery (Myeloma):			
	Leukocytes ($x10^9$/L)	5.3 (min-max: 2-8.7)	4.3 (min-max: 1.9-6.6)	0.14
	Hemoglobin (g/dL)	12.2 (min-max: 8.2-16.9)	12 (min-max: 8.9-14.5)	0.64
	Platelets ($x10^9$/L)	220.9 (min-max: 78-361)	258.9 (min-max: 138-369)	0.05
Manko et al., 2014[78]	Number of PB CD34+ cells (/μL)	62.0 (2-394)	47.5 (2-370)	-
Schmitt et al., 2013[81]	WBC count in peripheral blood (10^9/L)	50.8 (min-max: 29.9-64.6)	43.3 (min-max: 27.1-62.5)	0.27
Yafour et al., 2013[83]	WBC count in peripheral blood (10^9/L) (day 5)	35.5 (min-max: 26.6-65.4)	37.5 (min-max: 20.9-67.7)	0.96

Neutropenia / chemotherapy disturbance:

Author, year	(Primary) Outcome Measure(s)	Outcome Biosimilar	Outcome Reference Biological	P-value
Gascon et al., 2013[84]	Episode of febrile neutropenia	N=29 (2.2%)	No control group used	
	Incidence of neutropenia (grade 4)	N=104 (8.5%)	No control group used	-
	Chemotherapy regimens disturbance	Study I: 8/77 (10%), Study II: 27/307 (7%)	No control group used	-
Sagara et al., 2013[87]	Duration of neutropenia	2.2±1.5 days	No control group used	-
Salesi et al., 2012[88]	Incidence of febrile neutropenia	N=3 (6.25%)	No control group used	-
	Incidence of neutropenia (grade 4)	N=6 (12.5%)	No control group used	-
Verpoort & Mohler, 2012[90]	Incidence of neutropenia	N=1	N=1	-
	Chemotherapy dose reduction	N=5 (6.5%)	N=2 (8%)	-
	Chemotherapy dose discontinuation	N=2 (2.5%)	N=2 (8%)	-

Other outcome measures:

Author, year	(Primary) Outcome Measure(s)	Outcome Biosimilar	Outcome Reference Biological	P-value
Castelli et al., 2014[91]	Erythroid response (ER)	67% got ER	No control group used	-
	Number of transfusions needed.	62.5% got independent from transfusion, 29.1% did not respond to the biosimilar	No control group used	-
Kerkhofs et al., 2012[86]	Hb response	N=113 (79%)	No control group used	-
	Hb response rate (30,000 IU/week)	N=116 (81%)	No control group used	-
	Hb response rate (40,000 IU/week)	N=112 (78%)	No control group used	-
Remenyi et al., 2014[80]	Time to absolute neutrophil engraftment (days)	9 (min-max: 8-11)	No control group used	-
	Time to absolute leukocyte engraftment (days)	10 (min-max: 8-12)	No control group used	-
	Time to platelet engraftment (days)	10.5 (min-max: 7-19)	No control group used	-
Schmitt et al., 2014[89]	Differences in effectiveness	None reported	No control group used	-