POST-MARKETING SAFETY LEARNING FOR BIOLOGICALS

regulatory and clinical insights

Lotte A. Minnema

POST-MARKETING SAFETY LEARNING FOR BIOLOGICALS: REGULATORY AND CLINICAL INSIGHTS

Lotte Anne Minnema

Colophon

The studies presented in this thesis have been conducted under the umbrella of the Regulatory Science collaboration between the Dutch Medicines Evaluation Board (CBG-MEB) and the Utrecht Institute for Pharmaceutical Sciences (UIPS). The CBG-MEB is dedicated to ensure that licensed medicinal products during their whole life-cycle have a positive benefit-risk. This role requires intensive collaboration with academic and clinical partners in order to develop new assessment and decision-making methods, to engage with the clinic and to strengthen regulatory science. This PhD thesis aims to go beyond its scientific merits as such by delivering science, learning and insight to promote public health.

Printing of this thesis was kindly supported by the CBG-MEB and UIPS.

ISBN: 978-94-93197-65-7

Cover, layout, and printing: Off Page, Amsterdam

Cover illustration: Astibuag. *Weather forecast heat map. Atmosphere front motion. Warm and cold air masses visualization. Fluid motion*. 2021. Photograph. Shutterstock. Web. 26 April 2020.

POST-MARKETING SAFETY LEARNING FOR BIOLOGICALS: REGULATORY AND CLINICAL INSIGHTS

Verwerven van kennis over de veiligheid van biologicals na markttoelating: regulatoire en klinische inzichten

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 28 juni 2021 des middags te 2.15 uur

door

Lotte Anne Minnema
geboren op 27 december 1991
te Nijmegen

PROMOTOREN

Prof. dr. H.G.M. Leufkens

Prof. dr. A.C.G. Egberts

COPROMOTOREN

Dr. T.J. Giezen

Dr. H. Gardarsdottir

TABLE OF CONTENTS

1

GENERAL INTRODUCTION

Throughout history, biological drugs ensured medical breakthroughs that changed the prognosis of diseases that previously could not be controlled or were incurable. In the 1890s, serum therapies, derived from horse plasma, were introduced as a therapy for diphtheria. These were shown to achieve a cure rate of approximately 80%, thereby reducing the mortality rate of diphtheria (1). In the 1920s, animal-derived insulin was first used in diabetic children who would previously have died from diabetic ketoacidosis (2). Later, with the introduction of recombinant DNA technology in the 1980s, recombinant-produced insulin became available, which optimized the therapy of diabetic patients because it was less immunogenetic than the available insulin derived from animals (3, 4). In the 1990s, targeted therapies such as the TNF-α inhibitor infliximab changed the prognosis of chronic immune-mediated inflammatory diseases (5, 6). With the inhibition of TNF-α, the underlying pathophysiology of the disease was targeted. Thus, in patients with rheumatoid arthritis, radiographic progression was prevented and joint integrity preserved (6). More recently, in the 2010s, immune checkpoint inhibitors heralded a new era in cancer treatment, using the patient's immune system as facilitator of the treatment (7). These checkpoint inhibitors have changed the prognosis of multiple cancer types, including advanced melanoma. Where advanced melanoma was uncurable in the past, the prognosis has greatly improved, and a subset of patients treated with checkpoint inhibitors have even shown durable responses (8, 9). Recently, cell and gene therapies offering new treatment modalities have also started to revolutionize clinical practice (10). In addition, the vaccines for the prevention of Covid-19 are expected to be a breakthrough in the combat against this pandemic. Nowadays, biologicals represent approximately 30% of the newly launched active substances worldwide, and this is expected to increase in the coming years (11).

Over time, the definitions that have been applied in classifying drugs as biological drugs have also evolved. In past literature, biologicals were generally considered to be therapeutic products produced by modern biotechnological techniques. By applying this criterion, for example, serum therapies and insulin derived from animals would not be classified as biologicals. The definition has been amended with the introduction of diagnostic products and nucleic acid-based and cell therapies. Currently, the European Medicines Agency (EMA) states that a biological substance is a substance that is produced by or extracted from a biological source and that needs for its characterization and the determination of its quality a combination of physico-chemical-biological testing together with the production process and its control (12). With these definitions, biologicals are distinguished from small molecules, as small molecules are generally produced through chemical synthesis and have relatively simple structures that can be adequately characterized. Although biologicals are more complex than small molecules, differences in structural complexity exist within the group. Biologicals vary from relatively simple structures, such as insulin, to complex protein structures, such as monoclonal antibodies and blood coagulation factors.

To ensure that both the biologicals and small molecules can be safely and effectively applied in patient care, they must be regulated by an independent body. Influenced by many historical events, the regulatory system has evolved into the system that is in place today (Figure 1). Nowadays, companies that develop drugs are required to perform extensive research prior to marketing approval of these to ensure the pharmaceutical quality, efficacy, and safety of a drug. If, based on the data provided, the regulatory authorities consider that a consistent quality is demonstrated and

the benefits of a drug outweigh the risks in the treated population, the drug is approved. However, uncertainties about the safety and efficacy of a drug always remain at the time of marketing approval. The clinical trials that serve as the main generators of evidence regarding the efficacy and safety of a drug have their limitations as, often, strict patient eligibility criteria may be applied, a limited number of patients is included, and the duration of follow-up is limited (13). Given that clinical trials are not able to detect adverse events occurring rarely or with a long latency, the safety profile in particular should be further characterized when the drug is used in clinical practice. For this purpose, pharmacovigilance is in place, which is defined by the World Health Organization as "the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other medicine-related problem" (14). The establishment of the pharmacovigilance system was prompted by the thalidomide tragedy in the 1960s, which demonstrated that systematic collection of safety data when the drug is being used in clinical practice is vital (15). During the last decades, the regulatory system moved towards a more proactive risk management approach prompted by, among others, the withdrawal of rofecoxib and the suspension of rosiglitazone, leading up to the introduction of the new pharmacovigilance legislation in 2012.

The pharmacovigilance system has multiple tools available to continuously monitor the safety profile of drugs in the post-marketing phase. For all drugs, companies are required to collect all suspected adverse drug reactions that are reported by patients and healthcare professionals (16). These reports are also included in a European database (EudraVigilance) and in a global database (VigiBase), which is maintained by the Uppsala Monitoring Centre in Sweden. For drugs for which less information is available, for example, in the first years after approval, the reporting of suspected adverse drug reactions is enhanced by actively encouraging patients and healthcare professionals to report them. The reports of the suspected adverse drug reactions are evaluated by companies and regulatory authorities in order to assess whether it reflects new safety information. This includes a new association between the drug and an adverse event or a new aspect to a known association, such as the frequency or severity of an adverse event, which is then considered to be a safety signal. When there is a safety signal, this can be evaluated by the EMA's Pharmacovigilance Risk Assessment Committee (PRAC). A significant proportion of the safety signals discussed at the PRAC concerns biologicals (Figure 2).

Another pharmacovigilance tool for monitoring and evaluating the safety of a drug in the post-marketing phase is the periodic safety update report (PSUR). The PSUR for an individual drug lists all relevant safety information and includes, among others, information about spontaneous reports and safety results of ongoing studies and is periodically submitted to the regulatory authorities by the drug companies. Previous empirical work has shown that potential safety issues were identified in 83% of the PSURs for biologicals and that it was concluded that the product information should be updated in 37% of the PSURs (18). The general pharmacovigilance activities of detecting potential safety signals and submitting PSURs are complemented by activities tailored to further characterize specific safety issues. At the time of approval, the safety issues that should be further characterized in the post-marketing phase are described in the risk management plan (RMP) (19). The core of the RMP is the safety specification that lists the safety concerns for which a distinction is made in the "important identified risks," "important potential risks," and "missing information" and is

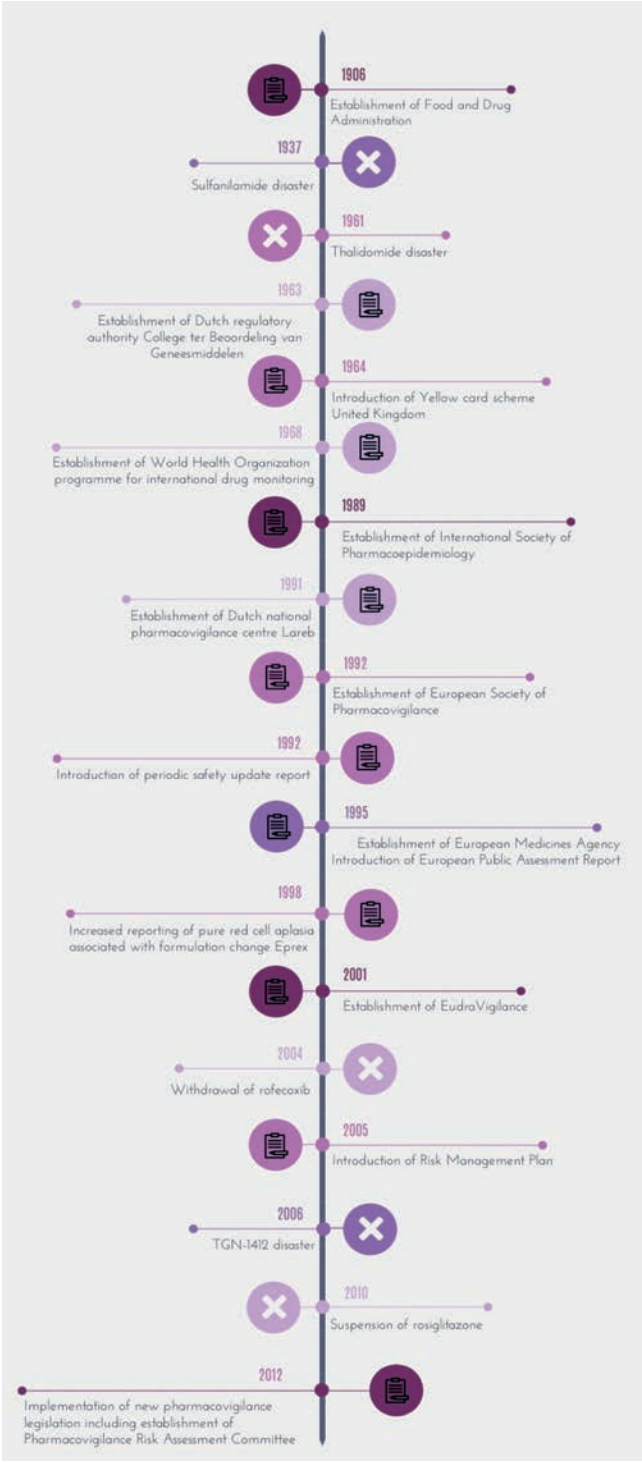


Figure 1. Key events in the evolution of regulatory pharmacovigilance

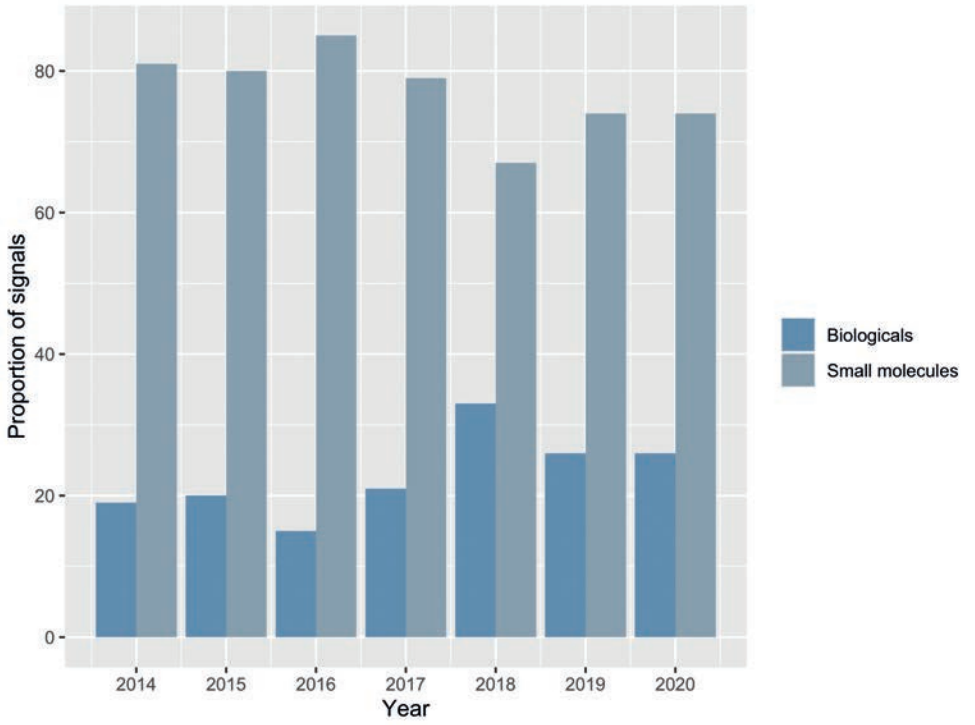


Figure 2. Proportion of safety signals discussed by the Pharmacovigilance Risk Assessment Committee for biologicals and small molecules in the period between 2014 and 2020 (Adapted from "List of safety signals discussed since September 2012") (17).

updated throughout the drug's life-cycle to reflect new safety information. Several studies have addressed the dynamics of the safety issues described in the RMP during the life-cycle of the drug. Vermeer et al. found that approximately 20% of the uncertainties ("important potential risks" or "missing information") described in the RMP were resolved within five years after approval (20). However, as approximately the same number of new uncertainties were added, the uncertainties remained stable over time. Duijnhoven et al. reported that uncertainties about the safety profile described in the RMP commonly concern a potential cancer risk, especially for biologicals (21). When the safety issues described in the RMP need specific follow-up, additional pharmacovigilance activities can be implemented. For example, in the pre-approval phase of brodalumab, a safety concern was raised regarding the occurrence of suicidal ideation and behavior (22). In addition to the measures taken to minimize the risk, the company was requested to perform an observational study to further characterize the risk of suicidal ideation and behavior. Furthermore, when specific populations (e.g., pregnant women) are not studied in the clinical trials and potential safety issues may be associated with use in these populations, further studies may be required. This was the case for infliximab, which is used for the treatment of chronic diseases prevalent in women with childbearing potential and for which the preclinical studies indicated a potential negative effect

for the infants (23). As in the clinical trials pregnant women were mostly excluded due to ethical reasons, limited information about the safety profile in these women was available. Therefore, in the post-marketing phase, the infants born to mothers exposed to infliximab were included in and followed over time in a registry study (23).

The pharmacovigilance system is challenged when dealing with more uncertainties at the time of approval. Specific regulatory pathways are in place that aim to provide patient access to drugs that address an unmet medical need. For these drugs, at the time of approval, less comprehensive data are required than for drugs that are approved through regular pathways, which inherently increases uncertainty about the safety and efficacy profile. As a consequence, evidence generation about safety of these drugs is shifted even more toward the post-marketing phase than for drugs approved through regular pathways. For example, blinatumomab was shown to be effective in adult patients with acute lymphoblastic leukemia and therefore to fulfill an unmet medical need (24). However, the clinical studies did not include a comparator arm, which limited the possibility of assessing the causality of the adverse events identified in them. Moreover, the safety was only studied in a total of 475 patients. Therefore, the company was, among other activities, obliged to study the safety profile in the post-marketing setting (24). Although post-marketing safety learning aims to minimize uncertainties, this can be challenging to achieve, in particular for drugs such as blinatumomab that are used for orphan diseases for which the number of patients receiving treatment is limited. The pharmacovigilance system is increasingly faced with this challenge as, within the last decade, the proportion of drugs indicated for orphan diseases has increased (25, 26). A significant proportion of the orphan drug approvals consists of biologicals; in the European Union, 39% of the orphan drug approvals in the period between 2007 and 2019 were biologicals (26).

In the past 20 years, specific challenges posed by biologicals to post-marketing safety learning have been addressed. These challenges include the difference in nature of adverse events from those known for small molecules, the complexity of the mechanism of action (including interference with the immune system), difficulties in classifying adverse events according to the established system, and the fact that the detection of adverse events is complicated when the symptoms of the adverse events mimic those of the treated disease (Table 1).

First, biologicals pose challenges to post-marketing safety learning because the nature of the adverse events of biologicals differs from that of small molecules. For biologicals, immunogenicity is more pronounced than for small molecules. Although the immunogenic potential of biologicals decreased with the introduction of recombinant DNA technologies and the humanization of monoclonal antibodies, immunogenicity is still a concern (27). The formation of the antidrug antibodies can induce different clinical effects. The effect of biologicals can be hampered directly by the formed binding or neutralizing antibodies or by altered pharmacokinetics. This is an important reason for the failure of hemophilia treatment with factor VIII products (28). The antibodies can also cross-react with endogenous factors, as is the case for the epoetins, for which prolonged treatment is associated with the formation of anti-erythropoietin antibodies resulting in pure red cell aplasia (29). Furthermore, a higher concentration of antibodies is associated with an increased

Table 1. Challenges related to post-marketing safety learning for biologicals including examples thereof.

Challenges related to post-marketing safety learning of biologicals	Examples
Difference in nature of adverse events from those known for small molecules	Occurrence of pure red cell aplasia following the cross-reaction of antibodies with endogenous erythropoietin during epoetin treatment
	Product characteristics affected antibody-dependent cell-mediated toxicity leading to reduced effectiveness of Herceptin
Complexity of the mechanism of action (including interference with the immune system)	Unexpected immune-related adverse events in patients treated with daclizumab since the mechanism of action was not fully known
	Difficulties to disentangle the association between brodalumab treatment and suicidal ideation and behavior due to limited knowledge of the influence of brodalumab on the central nervous system
Difficulties in classifying adverse events according to the established system	Immunogenicity and hypersensitivity reactions cannot be classified according to the typical distinction between type A and B adverse events
Detection of adverse events complicated when the symptoms of the adverse events mimic those of the treated disease	Early symptoms of the adverse events encephalitis and meningoencephalitis associated with daclizumab treatment mimicked the symptoms of multiple sclerosis relapse
	Misinterpretation of progressive multifocal leukoencephalopathy associated with natalizumab treatment as multiple sclerosis relapse

risk of infusion reactions (30). Many aspects influence the degree of antibody formation, including factors such as concomitant use of immunosuppressants and patient-related genetic factors (31-33). In addition, product-related factors, such as the glycosylation profile and the presence of impurities, play a role. Another product-related factor that is of influence is the manufacturing process, including the formulation of the product. The potential consequences of changes in the formulation are illustrated by the landmark example of epoetin-α. Following the replacement of human serum albumin by polysorbate 80 and glycine, an increase in the number of reports of pure red cell aplasia was observed (34, 35). Other changes in the manufacturing process can also result in variability of the product characteristics among batches of the same product as well as among the originator biological product and biosimilar. For trastuzumab, the biosimilar product Ontruzant showed a higher event-free survival compared with the reference product Herceptin (36). This difference could, however, be explained by differences in physicochemical and biological properties of the reference product in lots with expiry dates between August 2018 and December 2019 (36, 37). These differences affected the antibody-dependent cell-mediated toxicity and therefore the effectiveness of Herceptin. Given this variability of product characteristics caused by manufacturing changes, the EMA requires that for biologicals it should be ensured that the product and batch is identifiable (38). Previous empirical work has shown that the identification of the product in reported adverse events is adequate, whereas there is room for improvement for the identification of the batch (39). Recently, the importance of the identification of the batch

and product has been emphasized by the EMA and national authorities in the light of the Covid-19 vaccination campaigns (40). In addition to the occurrence of immunogenicity reactions, the nature of other adverse events also differs between biologicals and small molecules. For biologicals it was shown that, compared with small molecules, adverse events that were related to infections and infestations were reported more frequently, whereas psychiatric disorders and vascular disorders were reported less frequently (41). Another study illustrated that these differences were not attributable to differences in the indications for which the products are used, as biologicals and small molecules used for the same diseases also differed in the nature of the observed adverse events (42). For example, for biological immunosuppressants, the adverse events of neoplasms and infections and infestations occurred more frequently for biopharmaceuticals than for small-molecule immunosuppressants (20% vs 2%, 22% vs 9%, respectively) (42). Secondly, post-marketing safety learning is challenged by the complexity of the mechanism of action of biologicals. Although for biologicals it is known that the adverse events are often assigned to an exaggerated pharmacological response, the mechanism of action is not always fully elucidated. This is especially relevant for biologicals that exert their pharmacological effect through interference with the immune system. Given that the immune system is very complex and the knowledge of the specific pathways that are involved may be limited, the assessment of the association between the drug and the occurrence of the adverse event may be hampered. This is illustrated by the potential safety issue of suicidal ideation and behavior observed for brodalumab. In the clinical trials studying the efficacy and safety of brodalumab in patients with plaque psoriasis, several cases of completed suicide were observed (43). Brodalumab acts through the inhibition of the IL-17 receptor and acts therefore on the IL-17 axis for which the non-clinical data suggests that it can play a role in depression (22, 44). However, little is known about the effects of IL-17 in the human central nervous system. The effect of brodalumab on this system could therefore not be excluded, although a firm underlying mechanism could also not be established based on the limited information available (22). Similarly, the complexity of the influence of biologicals on the immune system was also seen for daclizumab, an antibody against CD-25 used for the treatment of relapsing forms of multiple sclerosis. Several years after the approval of daclizumab, immune-related events in the central nervous system (encephalitis and meningoencephalitis) were observed in clinical practice (45). At the time of approval, the company and regulatory authorities did not identify immune-related adverse events as a (potential) risk for daclizumab (46). However, in the clinical trials supporting its approval, there may have been suggestions for the occurrence of these adverse events in patients using this drug, as the incidence of immune-related adverse events was higher for daclizumab compared with placebo (45). As the mechanisms of action of daclizumab are not fully known, the mechanism through which it causes the immune-related adverse event was also not elucidated.

Thirdly, another challenge with biologicals is that the adverse events associated with biologicals may be difficult to classify according to the typical distinction between type A and type B adverse events. Traditionally, type A adverse events are considered to be related to the pharmacological effect of the drug, to be dose-dependent, and to occur frequently, whereas type B adverse events are

considered to be unexpected and unpredictable, are not dose-dependent, and are uncommon (47). For small-molecule drugs, type B events can include a variety of immunological reactions, including anaphylaxis, dermatitis, and vasculitis. For biologicals, however, these events are not unexpected and can be even related to the mechanism of action, such as in the case of the immune checkpoint inhibitors. Moreover, although no cases of pure red cell aplasia were reported in the clinical trials studying epoetin-α, it was expected that anti-drug antibody formation could occur. Therefore, alternative classification systems for biologicals have been described. For example, Lee and Kavanaugh differentiated between target-related and agent-related adverse events (48). Pichler et al. further specified the adverse events and proposed dividing the adverse events for biologicals (cytokines, antibodies, and fusion proteins) into the following five groups: high cytokine levels (type α); hypersensitivity because of an immune reaction against the biological agent (β); immune or cytokine imbalance syndromes (γ); symptoms due to cross–reactivity (δ); and symptoms not directly affecting the immune system (ε) (49). These classification systems can facilitate the further characterization of the safety profile of biologicals or, in clinical practice, facilitate the choice of the intervention following the occurrence of an adverse event. For this, however, the classification systems should be able to be used for, for example, signal detection purposes, which was found to be difficult for the system proposed by Pichler et al. (50).

Fourthly, the detection of adverse events for biologicals can be complicated when the symptoms of the adverse events mimic those of the treated disease, which brings challenges to post-marketing safety learning. In the daclizumab example, events of encephalitis and meningoencephalitis were observed in clinical practice. The early symptoms of these events include aphasia, confusion, and disorientation and were at first misinterpreted as worsening of the disease, as these are similar to the symptoms associated with multiple sclerosis relapse (45). The misinterpretation of adverse events was also seen in patients who developed progressive multifocal leukoencephalopathy (PML) when treated with natalizumab (51). When adverse events are not recognized as such by the patients and healthcare professionals in clinical practice, the characterization of the safety profile is hampered, as spontaneous reports are an important information source for identifying post-marketing safety issues, for small-molecule drugs as well as for biologicals (52).

THE ROLE OF SAFETY INFORMATION IN POST-MARKETING SAFETY LEARNING

Safety information can be provided through different information sources, which are, in the context of this thesis, divided into regulatory and clinical information sources. The regulatory information sources that are in place for all drugs include the European Public Assessment Report (EPAR) and product information. The EPAR describes the information supporting approval, including results of the preclinical and clinical studies and is updated throughout the drug's life-cycle when, for example, the indication of the drug is extended. The information described in the EPAR is translated into the product information that describes, among other things, the adverse events that are associated with the drug and the populations for which it is not safe to use the drug, and is also kept up to date throughout the drug's life-cycle. In its turn, the product information forms the basis of

the information provided in drug compendia used in clinical practice, such as *FASS* in Sweden and the *Farmacotherapeutisch Kompas* in the Netherlands (53, 54). In addition, the product information is used within the information systems used by healthcare professionals, including pharmacists and physicians. Besides the EPAR and product information that inform healthcare professionals about the safety profile of drugs, a Direct Healthcare Professional Communication is sent by the company, when new and important safety information becomes available, to directly inform healthcare professionals about this. Other information sources are the publications in peer-reviewed scientific journals. The safety results of the phase three clinical trials supporting the approval of new drugs are published, often in high-impact journals (55). Moreover, the results of post-marketing safety studies are published, either in peer-reviewed scientific journals or on the website of the European Network of Centers for Pharmacoepidemiology and Pharmacovigilance (56). These scientific publications, with other evidence-based information, provide the basis for the recommendations described in clinical guidelines. In addition to the safety information aimed at informing healthcare professionals, patients can also use a variety of information sources to gain knowledge about the safety profile of drugs. For drugs that are used by the patient, the package leaflet, as part of the product information, provides information about the adverse events that can be experienced. The package leaflet also forms the basis for specific webpages dedicated to promoting safe medication use by patients, such as *apotheek.nl* in the Netherlands.

These information sources play an important role in post-marketing safety learning. First, safety information can be aimed at patients and healthcare professionals to inform them about the safety profile of the drug in order to minimize the risks associated with the use of the drug. For example, at the time of the approval of the checkpoint inhibitors, there was limited experience in clinical practice with the associated immune-related adverse events. Therefore, extensive measures were implemented to inform both healthcare professionals and patients about the symptoms of the immune-related adverse events in order to identify and treat them early (57, 58). In the post-marketing phase, healthcare professionals gained experience with the use of the checkpoint inhibitors and their adverse events, and this eventually led to the discontinuation of the extensive measures (59, 60). Secondly, information can be provided about the populations for which it is not safe to use the drug or for which limited information about the safety profile was gained through the clinical trials. In the daclizumab example, in the clinical trials and post-marketing phase, it was shown that daclizumab was associated with unpredictable and potentially fatal liver injury (46). Therefore, the safety information states that daclizumab should not be used in patients with pre-existing liver disease. With the provision of safety information that describes the known adverse events, the generation of knowledge about new adverse events for the same drug or other drugs is also facilitated. Previously, it was shown that the identification of PML associated with natalizumab may have contributed to the identification of rituximab-associated PML, as the number of PML reports for rituximab increased after the safety information regarding natalizumab-associated PML was issued (61). Furthermore, healthcare professionals noticed the occurrence of the unexpected immune-mediated adverse events in patients using daclizumab, after which a thorough evaluation of the benefit–risk balance was initiated. In addition to informing healthcare professionals and patients about the safety profile of the drugs, the information provided through the different sources can

facilitate regulatory science. By providing information about the rationale of the decisions made by the regulatory authorities and keeping track of the implemented changes, regulatory decision-making can be studied.

Although all stakeholders benefit from adequate safety information, the intended purpose of the type of information source determines the safety information that is presented, which can differ among stakeholders. Moreover, the format of these information sources varies. For example, the package leaflet describes all adverse events that can occur during treatment, whereas healthcare professionals may discuss only the most frequently occurring or serious adverse events. Moreover, scientific publications are limited by the word count required by the journals, whereas the document size of the product information and clinical guidelines is not limited. Given these aspects, the safety information provided within the different documents may differ. There is, however, limited information available about the differences between the safety information provided through regulatory and clinical sources, which is addressed in this thesis.

THESIS AIM

Previous research and PhD theses from our group have focused on post-marketing safety learning for biologicals and have assessed the characterization of the safety profile and safety assessment of biologicals in the post-marketing phase and the regulatory tools that are in place for this purpose. A variety of safety-related regulatory actions and regulatory activities for biologicals have been studied and, if applicable, compared with other drug classes (18, 41, 42, 52, 62, 63). Furthermore, studies have addressed specific adverse events and the dynamics of safety learning on both unexpected adverse events and uncertainties regarding the safety profile at approval (20, 21, 61, 64).

Post-marketing safety learning for biologicals is continuously evolving, and, with the continuous introduction of biologicals with new mechanisms of action, additional challenges are introduced into the system. Moreover, specific aspects of the dynamics surrounding post-marketing safety learning as well as the link between information from regulatory and clinical sources have not been explored to date.

This thesis provides further insights in post-marketing safety learning for biologicals build on the established knowledge on this topic. We specifically focus on the characterization of specific adverse events, dynamics in post-marketing safety learning, and the comparison between safety information from regulatory and clinical sources.

THESIS OUTLINE

This thesis includes six studies divided over three chapters, followed by a general discussion.

Chapter 2 focuses on the characterization of specific safety issues for biologicals. In *Chapter 2.1*, we explore the association between use of monoclonal antibodies and depression and suicidal ideation and behavior using spontaneously reported adverse events. The study serves as a first step in studying this potential association for the group of monoclonal antibodies and explores the potential influence of their immunomodulating properties. In *Chapter 2.2*, we study

the incidence, longitudinal pattern, and potential risk factors of thyroid disorders in a cohort of patients treated with PD-1/PD-L1 inhibitors.

Chapter 3 addresses the dynamics in post-marketing safety learning during the drug's life-cycle and across biological products. In *Chapter 3.1*, we describe the post-marketing changes in the dosing information of biologicals. For small molecules, dosing changes occur frequently and are often safety related, whereas to date the frequency and nature of dosing changes have not been characterized for biologicals. Within the current study, we examined the number of dosing changes and the underlying rationale for the changes in dosing information. *Chapter 3.2* evaluates whether the overlap in adverse events described in the product information of TNF-α inhibitors is achieved during the drug's life-cycle. In addition, factors associated with the overlap of the described adverse events were studied.

Chapter 4 describes the comparison between safety information from regulatory and clinical information sources. In *Chapter 4.1*, we focus on biologicals used for treating patients with multiple sclerosis and compare the information about adverse events originating from regulators with that originating from the scientific community. The number and types of adverse events and the attention given to the adverse events were compared in order to study the potential consequences of the differences between the documents for optimal clinical decision making. In *Chapter 4.2*, we assess the methods used to identify and classify thyroid disorders in clinical trials and observational studies.

Finally, in the general discussion (**Chapter 5**), the results of the previous chapters are discussed in a broader context and recommendations for future clinical and regulatory practice are provided.

REFERENCES

1. Kaufmann SHE. Emil von Behring: translational medicine at the dawn of immunology. Nature Reviews Immunology. 2017;17(6):341-3.
2. Vecchio I, Tornali C, Bragazzi NL, Martini M. The Discovery of Insulin: An Important Milestone in the History of Medicine. Front Endocrinol (Lausanne). 2018;9:613.
3. Johnson IS. Human insulin from recombinant DNA technology. Science. 1983;219(4585):632-7.
4. Schernthaner G. Immunogenicity and allergenic potential of animal and human insulins. Diabetes Care. 1993;16 Suppl 3:155-65.
5. Maini R, St Clair EW, Breedveld F, Furst D, Kalden J, Weisman M, et al. Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. ATTRACT Study Group. Lancet. 1999;354(9194):1932-9.
6. Melsheimer R, Geldhof A, Apaolaza I, Schaible T. Remicade((R)) (infliximab): 20 years of contributions to science and medicine. Biologics. 2019;13:139-78.
7. Robert C. A decade of immune-checkpoint inhibitors in cancer therapy. Nature Communications. 2020;11(1):3801.
8. Luke JJ, Flaherty KT, Ribas A, Long GV. Targeted agents and immunotherapies: optimizing outcomes in melanoma. Nature Reviews Clinical Oncology. 2017;14(8):463-82.
9. Hamid O, Robert C, Daud A, Hodi FS, Hwu WJ, Kefford R, et al. Five-year survival outcomes for patients with advanced melanoma treated with pembrolizumab in KEYNOTE-001. Annals of Oncology. 2019;30(4):582-8.
10. Fischbach MA, Bluestone JA, Lim WA. Cell-based therapeutics: the next pillar of medicine. Sci Transl Med. 2013;5(179):179ps7.
11. IQVIA Institute for Human Data Science. The Global Use of Medicine in 2019 and Outlook to 2023. https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/the-global-use-of-medicine-in-2019-and-outlook-to-2023.pdf?_=1606729931964. Accessed 12/10/2020.
12. European Commission. Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32001L0083&from=en. Accessed 12/10/2020.
13. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet. 2005;365(9453):82-93.
14. World Health Organization. The importance of pharmacovigilance. https://apps.who.int/iris/handle/10665/42493. Accessed 1/11/2020.
15. McBride WG. THALIDOMIDE AND CONGENITAL ABNORMALITIES. The Lancet. 1961;278(7216):1358.
16. European Commission. Regulation (EC) No 726/2004 of the European parliament and of the council of March 2004 laying down community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European medicines agency. https://ec-europa-eu.proxy.library.uu.nl/health/sites/health/files/files/eudralex/vol-1/reg_2004_726/reg_2004_726_en.pdf. Accessed 1/11/2020.
17. European Medicines Agency. List of safety signals discussed since September 2012. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/signal-management/prac-recommendations-safety-signals#list-of-safety-signals-discussed-since-september-2012-section. Accessed 15/11/2020.
18. Ebbers HC, Mantel-Teeuwisse AK, Sayed-Tabatabaei FA, Moors EH, Schellekens H, Leufkens HG. The role of Periodic Safety Update Reports in the safety management of biopharmaceuticals. Eur J Clin Pharmacol. 2013;69(2):217-26.

19. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP): Module V – Risk management systems (Rev 2) https://www-ema-europa-eu.proxy.library.uu.nl/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-module-v-risk-management-systems-rev-2_en.pdf. Accessed 15/11/2020.
20. Vermeer NS, Duijnhoven RG, Straus SM, Mantel-Teeuwisse AK, Arlett PR, Egberts AC, et al. Risk management plans as a tool for proactive pharmacovigilance: a cohort study of newly approved drugs in Europe. Clin Pharmacol Ther. 2014;96(6):723-31.
21. Duijnhoven RG, Straus SM, Raine JM, de Boer A, Hoes AW, De Bruin ML. Number of patients studied prior to approval of new medicines: a database analysis. PLoS Med. 2013;10(3):e1001407.
22. European Medicines Agency. Kyntheum: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/kyntheum-epar-public-assessment-report_en.pdf. Accessed 25/7/2020.
23. European Medicines Agency. Remicade: EPAR - Scientific Discussion. https://www.ema.europa.eu/en/documents/scientific-discussion/remicade-epar-scientific-discussion_en.pdf. Accessed 15/11/2020.
24. European Medicines Agency. Blincyto: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/blincyto-epar-public-assessment-report_en.pdf. Accessed 15/11/2020.
25. Miller KL, Lanthier M. Investigating the landscape of US orphan product approvals. Orphanet Journal of Rare Diseases. 2018;13(1):183.
26. IQVIA Institute for Human Data Science. The Prospects for Biosimilars of Orphan Drugs in Europe. https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/the-prospects-for-biosimilars-of-orphan-drugs-in-europe.pdf?_=1606758208060. Accessed 12/10/2020.
27. Kuriakose A, Chirmule N, Nair P. Immunogenicity of Biotherapeutics: Causes and Association with Posttranslational Modifications. J Immunol Res. 2016;2016:1298473.
28. Leissinger C, Cooper DL, Solem CT. Assessing the impact of age, race, ethnicity and inhibitor status on functional limitations of patients with severe and moderately severe haemophilia A. Haemophilia. 2011;17(6):884-9.
29. Macdougall IC, Roger SD, de Francisco A, Goldsmith DJ, Schellekens H, Ebbers H, et al. Antibody-mediated pure red cell aplasia in chronic kidney disease patients receiving erythropoiesis-stimulating agents: new insights. Kidney Int. 2012;81(8):727-32.
30. Baert F, Noman M, Vermeire S, Van Assche G, G DH, Carbonez A, et al. Influence of immunogenicity on the long-term efficacy of infliximab in Crohn's disease. N Engl J Med. 2003;348(7):601-8.
31. Krieckaert CL, Bartelds GM, Lems WF, Wolbink GJ. The effect of immunomodulators on the immunogenicity of TNF-blocking therapeutic monoclonal antibodies: a review. Arthritis Res Ther. 2010;12(5):217.
32. Billiet T, Vande Casteele N, Van Stappen T, Princen F, Singh S, Gils A, et al. Immunogenicity to infliximab is associated with HLA-DRB1. Gut. 2015;64(8):1344-5.
33. Bartelds GM, de Groot E, Nurmohamed MT, Hart MH, van Eede PH, Wijbrandts CA, et al. Surprising negative association between IgG1 allotype disparity and anti-adalimumab formation: a cohort study. Arthritis Res Ther. 2010;12(6):R221.
34. Boven K, Stryker S, Knight J, Thomas A, van Regenmortel M, Kemeny DM, et al. The increased incidence of pure red cell aplasia with an Eprex formulation in uncoated rubber stopper syringes. Kidney Int. 2005;67(6):2346-53.
35. Schellekens H, Jiskoot W. Erythropoietin-Associated PRCA: Still an Unsolved Mystery. J Immunotoxicol. 2006;3(3):123-30.
36. Pivot X, Pegram M, Cortes J, Lüftner D, Lyman GH, Curigliano G, et al. Three-year follow-up from a phase 3 study of SB3 (a trastuzumab biosimilar) versus reference trastuzumab in the neoadjuvant setting for human epidermal growth factor receptor 2–positive breast cancer. European Journal of Cancer. 2019;120:1-9.

37. Kim S, Song J, Park S, Ham S, Paek K, Kang M, et al. Drifts in ADCC-related quality attributes of Herceptin®: Impact on development of a trastuzumab biosimilar. mAbs. 2017;9(4):704-14.
38. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP). Product- or Population-Specific Considerations II: Biological medicinal products. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-product-population-specific-considerations-ii_en-0.pdf. Accessed 16/12/2020.
39. Vermeer NS, Giezen TJ, Zastavnik S, Wolff-Holz E, Hidalgo-Simon A. Identifiability of Biologicals in Adverse Drug Reaction Reports Received From European Clinical Practice. Clin Pharmacol Ther. 2019;105(4):962-9.
40. European Medicines Agency. Pharmacovigilance Plan of the EU Regulatory Network for COVID-19 Vaccines. https://www.ema.europa.eu/en/documents/other/pharmacovigilance-plan-eu-regulatory-network-covid-19-vaccines_en.pdf. Accessed 16/12/2020.
41. Giezen TJ, Mantel-Teeuwisse AK, Meyboom RH, Straus SM, Leufkens HG, Egberts TC. Mapping the safety profile of biologicals: a disproportionality analysis using the WHO adverse drug reaction database, VigiBase. Drug Saf. 2010;33(10):865-78.
42. Ebbers HC, Al-Temimi E, Moors EH, Mantel-Teeuwisse AK, Schellekens H, Leufkens HG. Differences between post-authorization adverse drug reactions of biopharmaceuticals and small molecules. BioDrugs. 2013;27(2):167-74.
43. Schmidt C. Suicidal thoughts end Amgen's blockbuster aspirations for psoriasis drug. Nat Biotechnol. 2015;33(9):894-5.
44. Beurel E, Lowell JA. Th17 cells in depression. Brain Behav Immun. 2018;69:28-34.
45. European Medicines Agency. Zinbryta Article-20 referral - PRAC assessment report. https://www.ema.europa.eu/en/documents/referral/zinbryta-article-20-referral-prac-assessment-report_en-0.pdf. Accessed 25/7/2020.
46. European Medicines Agency. Zinbryta: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/zinbryta-epar-public-assessment-report_en.pdf. Accessed 25/7/2020.
47. Meyboom RH, Lindquist M, Egberts AC. An ABC of drug-related problems. Drug Saf. 2000;22(6):415-23.
48. Lee SJ, Kavanaugh A. Adverse reactions to biologic agents: Focus on autoimmune disease therapies. Journal of Allergy and Clinical Immunology. 2005;116(4):900-5.
49. Pichler WJ. Adverse side-effects to biological agents. Allergy. 2006;61(8):912-20.
50. Giezen TJ. Risk Management of Biologicals: A regulatory and Clinical Perspective: Utrecht University; 2011.
51. European Medicines Agency. Tysabri: EPAR - Scientific discussion. https://www.ema.europa.eu/en/documents/scientific-discussion/tysabri-epar-scientific-discussion_en.pdf. Accessed 17/01/2021.
52. Ebbers HC, Mantel-Teeuwisse AK, Moors EH, Sayed Tabatabaei FA, Schellekens H, Leufkens HG. A cohort study exploring determinants of safety-related regulatory actions for biopharmaceuticals. Drug Saf. 2012;35(5):417-27.
53. Swedish Association of the Pharmaceutical Industry. FASS Allmänhet. https://www.fass.se/LIF/startpage. Accessed 17/01/2021.
54. Zorginstituut Nederland. Farmacotherapeutisch Kompas. https://www.farmacotherapeutischkompas.nl/. Accessed 17/01/2021.
55. Unger JM, Barlow WE, Ramsey SD, LeBlanc M, Blanke CD, Hershman DL. The Scientific Impact of Positive and Negative Phase 3 Cancer Clinical Trials. JAMA Oncol. 2016;2(7):875-81.
56. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). The European Union electronic Register of Post-Authorisation Studies http://www.encepp.eu/encepp/studiesDatabase.jsp. Accessed 2/2/2021.

57. European Medicines Agency. Keytruda: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/keytruda-epar-public-assessment-report_en.pdf. Accessed 25/7/2020.

58. European Medicines Agency. Opdivo: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/opdivo-epar-public-assessment-report_en.pdf. Accessed 25/7/2020.

59. European Medicines Agency. Pharmacovigilance Risk Assessment Committee (PRAC): Minutes of the meeting on 12-15 March 2019. https://www.ema.europa.eu/en/documents/minutes/minutes-prac-meeting-12-15-march-2019_en.pdf. Accessed 15/11/2020.

60. European Medicines Agency. Pharmacovigilance Risk Assessment Committee (PRAC): Minutes of meeting on 10 – 13 February 2020. https://www.ema.europa.eu/en/documents/minutes/minutes-prac-meeting-10-13-february-2020_en.pdf. Accessed 15/11/2020.

61. Vermeer NS, Straus SM, Mantel-Teeuwisse AK, Hidalgo-Simon A, Egberts AC, Leufkens HG, et al. Drug-induced progressive multifocal leukoencephalopathy: Lessons learned from contrasting natalizumab and rituximab. Clin Pharmacol Ther. 2015;98(5):542-50.

62. Giezen TJ, Mantel-Teeuwisse AK, Straus SM, Schellekens H, Leufkens HG, Egberts AC. Safety-related regulatory actions for biologicals approved in the United States and the European Union. Jama. 2008;300(16):1887-96.

63. Giezen TJ, Mantel-Teeuwisse AK, Straus SM, Egberts TC, Blackburn S, Persson I, et al. Evaluation of post-authorization safety studies in the first cohort of EU Risk Management Plans at time of regulatory approval. Drug Saf. 2009;32(12):1175-87.

64. Giezen TJ, Mantel-Teeuwisse AK, ten Berg MJ, Straus SM, Leufkens HG, van Solinge WW, et al. Rituximab-induced thrombocytopenia: a cohort study. Eur J Haematol. 2012;89(3):256-66.

2

CHARACTERIZATION OF SPECIFIC SAFETY ISSUES

2.1

EXPLORING THE ASSOCIATION BETWEEN MONOCLONAL ANTIBODIES AND DEPRESSION AND SUICIDAL IDEATION AND BEHAVIOR: A VIGIBASE STUDY

Lotte A. Minnema, Thijs J. Giezen, Patrick C. Souverein,
Toine C.G. Egberts, Hubert G.M. Leufkens, Helga Gardarsdottir

Drug Saf. 2019 Jul;42(7):887-895

ABSTRACT

Introduction: Several monoclonal antibodies (mAbs) have been linked to neuropsychiatric adverse effects in patients, including depression and suicidal ideation and behavior.

Aim: The aim of this study was to quantify and characterize spontaneously reported adverse drug reactions (ADRs) of depression and suicidal ideation and behavior related to mAb users, and to explore a possible association with their mechanism of action.

Methods: We included mAb ADRs that were reported in VigiBase, and identified those related to depression and suicidal ideation and behavior. Reporting odds ratios (RORs) were estimated for each mAb (bevacizumab as the reference) and according to their influence on the immune system (not directly targeting [reference], stimulating, or suppressing). Those suppressing the immune system were further divided into their intended indication (auto-immune diseases, cancer).

Results: Overall, 2,924,319 ADRs for 44 mAbs were included; 9455 ADRs were related to depression and 1770 were related to suicidal ideation and behavior. The association was strongest for natalizumab and belimumab, both for depression (ROR 5.7, 95% confidence interval [CI] 5.0–6.4; and ROR 5.1, 95% CI 4.2–6.2) and suicidal ideation and behavior (ROR 12.0, 95% CI 7.9–18.3; and ROR 20.2, 95% CI 12.4–33.0). Those suppressing the immune system showed higher ROR, i.e. 1.9 (95% CI 1.8–2.0) for depression and 3.6 (95% CI 3.0–4.4) for suicidal ideation and behavior. This finding was only seen for mAbs used for treating autoimmune diseases.

Conclusion: Depression and suicidal ideation and behavior are seen in patients using mAbs, particularly mAbs used for treating autoimmune diseases that suppress the immune system. For interpretation of these data, the indications for use and other characteristics require further consideration.

INTRODUCTION

In May 2015, the phase III clinical trials investigating the efficacy and safety of brodalumab, a monoclonal antibody (mAb) against the interleukin (IL)-17 receptor, in patients with psoriasis were terminated early by the pharmaceutical company (1). The trigger for this decision was six reports of completed suicide as adverse events in the approximately 5000 patients treated with brodalumab in the clinical trial program (2), although the evaluation thereof by both the pharmaceutical company and regulatory authorities concluded that a causal relation between the use of brodalumab and suicidal ideation and behavior was unlikely (1, 2). In May 2017, brodalumab was authorized in the European Union based on the assessment of the regulatory authorities that the efficacy of brodalumab outweighs the risks, including the potential risk of suicidal ideation and behavior (2). Despite this conclusion, a warning was included in the product information to carefully weigh the risks and benefits of treatment with brodalumab for patients with a history of depression and/or suicidal ideations, and for patients who develop these symptoms during treatment. Furthermore, a post-authorization safety study was required with a focus on serious events of, among others, suicidal ideation and behavior. To date, a mechanism through which brodalumab may cause such events is not known. Research in animal studies has shown that IL-17, which is blocked by brodalumab, can influence neurological function and therefore modulate behavior (3-5); however, this has not been studied in humans.

Depression and suicidal ideation and behavior have also been reported in clinical trials of other mAbs, such as belimumab, which is used to treat patients with systemic lupus erythematosus and targets against B-cell activating factor (6). In addition, the mAbs infliximab, adalimumab, and natalizumab have, in case reports, been associated with suicidal ideation and behavior (7-11).

The aforementioned mAbs exert their intended effects through targeting different (anti-) inflammatory factors. During the past decades, the link between psychiatric disorders and autoimmune disorders has been extensively discussed (12-14). This link is considered to be partially explained by the influence of inflammatory factors on the brain. Different meta-analyses have evaluated the contribution of inflammatory factors in the pathophysiology of major depression and suicidal ideation and behavior (15-20). These indicate that alterations in both pro-inflammatory and anti-inflammatory factors are linked to psychiatric disorders; however, mAbs that have a mechanism of action not directly targeting the immune system have also been linked to psychiatric disorders. For example, approximately 1–10% of patients treated with trastuzumab, a human epidermal growth factor receptor 2 inhibitor used for the treatment of cancer, develop depression (21). It should be noted that the estimated prevalence of major depression among patients with both cancer and autoimmune diseases exceed the estimated prevalence in the general population (22, 23). It is therefore challenging to differentiate between the underlying disease and the effect of treatment.

To date, no studies have, to our knowledge, evaluated the potential risk of depression and suicidal ideation and behavior for the group of mAbs as a whole. Therefore, the first aim of this study was to quantify and characterize spontaneously reported adverse drug reactions (ADRs) related to depression and suicidal ideation and behavior for mAbs. In addition, the association between the mechanism of action of the mAb and spontaneously reported ADRs of depression and suicidal ideation and behavior will be explored.

METHODS

Setting and data source

VigiBase, the World Health Organization global individual case safety report (ICSR) database that is maintained by the Uppsala Monitoring Centre, was used as the data source for this study (24). As of December 2017, over 16 million case reports of ADRs have been submitted since the start of data collection in 1968. The ICSRs are first reported by healthcare professionals and patients to more than 120 national pharmacovigilance centers, and then transferred to VigiBase. ICSRs contain information regarding patient characteristics, suspected drugs, ADRs, and additional information relevant to the report, such as the reporter type, reporting year, and reporting region (24). Completeness of the ICSRs is variable. ADRs are classified according to the Medical Dictionary for Regulatory Activities (MedDRA®), and suspected drugs are classified according to the Anatomical Therapeutic Chemical (ATC) classification.

All reports included in VigiBase until December 2017 in which an mAb was the suspected drug were identified. Only mAbs that had been authorized by the European Medicines Agency (EMA) and/or the US FDA for 3 or more years as of December 2017 were included, accounting for a representative reflection of the ADRs that are reported. Information on the regulatory status was retrieved from the publicly available information on the FDA (http://www.fda.gov) and EMA (http://www.ema.europ a.eu) websites.

Outcome

Spontaneously reported ADRs related to depression and suicidal ideation and behavior were identified using the Standardized MedDRA® Query (SMQ) Depression and suicide/self-injury (narrow) [MedDRA® version 20.1]. SMQs are validated and maintained by the Maintenance and Support Services Organization and updated with each version of MedDRA® (25). The SMQ 'Depression and suicide/self-injury (narrow)' contains 36 preferred terms. A distinction is made in preferred terms related to depression (n =24) and suicide/self-injury (n =12).

Exposure

The association between the mAb and spontaneously reported ADRs of depression and suicidal ideation and behavior was defined in two ways. First, exposure was defined for the mAbs individually using bevacizumab as the reference. Second, mAbs were grouped by their influence on the immune system (not directly targeting the immune system, suppressing or stimulating the immune system) (see supplementary material). MAbs suppressing the immune system were further stratified according to their intended indication based on their ATC code (autoimmune diseases or cancer). Information on the influence of mAbs on the immune system and their intended indication was retrieved from the product information publicly available on the FDA and/or EMA websites.

Data analysis

Descriptive statistics were used to quantify the reported ADRs at the level of the drug–ADR pair. ADRs were stratified by sex (male, female), age (< 18 years, 18–44 years, 45–64 years, and ≥ 65 years), reporting year (from the first reporting year divided into periods of 5 years), reporting region (Africa, the Americas, Southeast Asia, Europe, Eastern Mediterranean, and Western Pacific), and reporter type (healthcare professional, consumer, other). The proportions of ADRs for depression and suicidal ideation and behavior were calculated by dividing the number of depression or suicidal ideation and behavior ADRs by the total number of reported ADRs within the stratum.

The strength of the association between mAbs and neuropsychiatric effects was expressed as the reporting odds ratio (ROR) with 95% confidence intervals (CIs) (26, 27). RORs were estimated for the ADRs related to depression and suicidal ideation and behavior separately. For analysis of the individual mAbs, bevacizumab was used as the reference as it does not directly target the immune system and has been widely used for years, and was therefore considered to have an established safety profile. For the analysis of the mAbs grouped by their influence on the immune system, mAbs not directly targeting the immune system were used as the reference. In addition, for the subgroup analysis stratified by the intended indication, mAbs not directly targeting the immune system were used as the reference.

All data were analyzed using SPSS for Windows, version 24.0 (IBM Corporation, Armonk, NY, USA).

RESULTS

In VigiBase, 3,048,884 ADRs were identified for 139 different mAbs as suspected drugs; 124,565 ADRs for 95 mAbs were excluded because they were reported for mAbs that had not been authorized by the FDA and/or EMA for 3 or more years as of December 2017. The study population therefore comprised 44 mAbs (active substances), for which a total of 1,048,576 ICSRs were filed, representing 2,924,319 drug–ADR pairs, of which 9455 (0.32%) were related to depression and 1770 (0.06%) were related to suicidal ideation and behavior.

The proportion of suicidal ideation and behavior ADRs was comparable between men and women (0.07% and 0.06%, respectively), whereas the proportion of depression ADRs was higher in women compared with men (0.35% and 0.27%, respectively) (Table 1). The highest proportion of depression and suicidal ideation and behavior ADRs was observed in the age range between 18 and 64 years, at 0.35% and 0.08%, respectively (Table 1). The highest proportion of depression ADRs originated from the Americas region (0.36%), followed by the European region (0.17%). No regional differences were seen in the proportion of suicidal ideation and behavior ADRs. The proportion of ADRs involving depression increased over time, from 0.11% in the period between 2000 and 2004, to 0.28% in the period between 2015 and 2017, whereas the frequency of suicidal ideation and behavior remained at approximately 0.06%. Among consumer reports, the proportion of depression ADRs was higher (0.44%) than among reports of healthcare professionals (0.23%), whereas for suicidal ideation and behavior, the proportion was comparable (0.05% and 0.07%).

For seven mAbs, no ADRs were reported for depression, and for 17 mAbs, no suicidal ideation and behavior ADRs were reported; therefore, the ROR could not be estimated for these mAbs.

Table 1. Characteristics of depression, suicidal ideation and behavior, and all adverse drug reactions (ADRs) reported for the monoclonal antibodies.

		Depression ADRs (%)	Suicidal ideation and behavior ADRs (%)
Total (n=2,924,319)		9,455 (0.32)	1,770 (0.06)
Sex	Male (n=889,618)	2,423 (0.27)	638 (0.07)
	Female (n=1,896,884)	6,669 (0.35)	1,072 (0.06)
	Unknown (n=137,817)	363 (0.26)	60 (0.04)
Age	<18 years (n=62,164)	97 (0.16)	61 (0.10)
	18 – 44 years (n=516,973)	1,830 (0.35)	515 (0.10)
	45 – 64 years (n=838,942)	2,822 (0.34)	525 (0.06)
	≥65 years (n=481,932)	868 (0.18)	98 (0.02)
	Unknown (n=1,024,308)	3,838 (0.37)	571 (0.06)
Region	African region (n=3,438)	7 (0.20)	1 (0.03)
	Region of the Americas (n=2,433,422)	8,648 (0.36)	1,523 (0.06)
	South-east Asia region (n=4,372)	2 (0.05)	3 (0.07)
	European region (n=376,078)	640 (0.17)	196 (0.05)
	Eastern Mediterranean region (n=3,252)	5 (0.15)	1 (0.03)
	Western Pacific region (n=103,757)	153 (0.15)	46 (0.04)
Reporting year	1995 – 1999 (n=1,367)	0 (0)	0 (0)
	2000 – 2004 (n=52,267)	57 (0.11)	5 (0.01)
	2005 – 2009 (n=284,580)	659 (0.23)	183 (0.06)
	2010 – 2014 (n=1,243,659)	4,928 (0.40)	782 (0.06)
	2015 – 2017 (n=1,342,446)	3,811 (0.28)	800 (0.06)
Reporter	Health care professional (n=1,413,828)	3,221 (0.23)	991 (0.07)
	Consumer (n=1,285,732)	5,600 (0.44)	633 (0.05)
	Other (n=29,983)	110 (0.37)	32 (0.11)
	Unknown (n=194,776)	524 (0.27)	114 (0.06)

For depression, the association (relative to bevacizumab) was strongest for natalizumab (ROR 5.7, 95% CI 5.0–6.4), followed by belimumab (ROR 5.1, 95% CI 4.2–6.2) (Fig. 1). Furthermore, RORs were calculated for different groups of mAbs based on their influence on the immune system (Fig. 2). For depression, the association was strongest for mAbs suppressing the immune system (ROR 1.9, 95% CI 1.8–2.0) when compared with mAbs not directly targeting the immune system. The mAbs suppressing the immune system were further stratified according to their intended indication. The results show that the association with depression for mAbs used for the treatment of autoimmune diseases was stronger compared with mAbs not directly targeting the immune system (ROR 1.96, 95% CI 1.84–2.10), whereas for mAbs used for the treatment of cancer, this difference was not seen (ROR 0.9, 95% CI 0.8–1.0). The mAbs stimulating the immune system were all indicated for the treatment of cancer and hence were not further stratified.

For suicidal ideation and behavior the ROR (relative to bevacizumab) was highest for belimumab (20.2, 95% CI 12.4–33.0) followed by natalizumab (12.0, 95% CI 7.9–18.3) (Fig. 3). When grouping the mAbs based on their influence on the immune system, the ROR for suicidal ideation and behavior was highest for mAbs suppressing the immune system (3.6, 95% CI 3.0–4.4) compared with mAbs not directly targeting the immune system (Fig. 4). When stratifying the mAbs suppressing

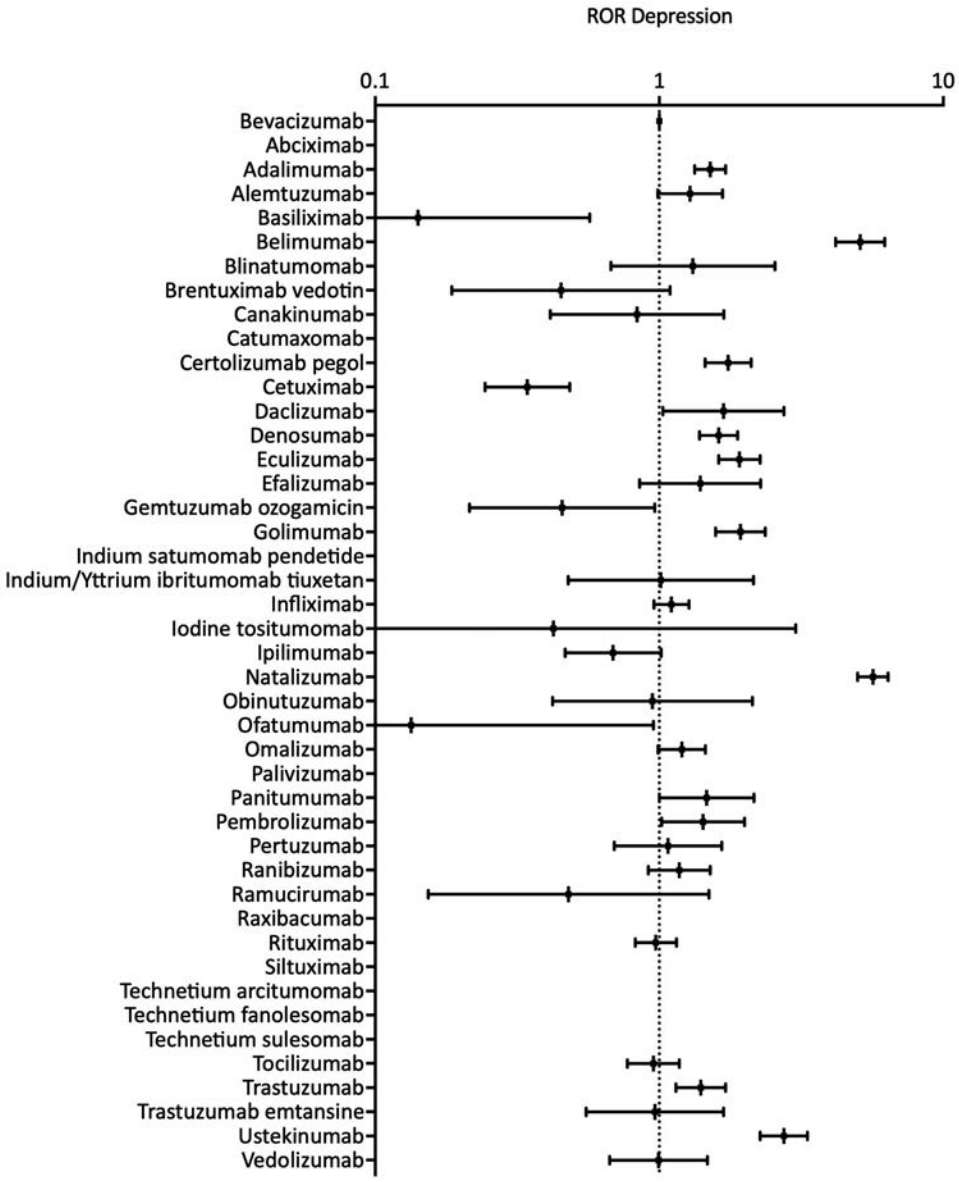


Figure 1. Reporting odds ratios (RORs) relative to bevacizumab of depression adverse drug reactions (ADRs) for the monoclonal antibodies separately.

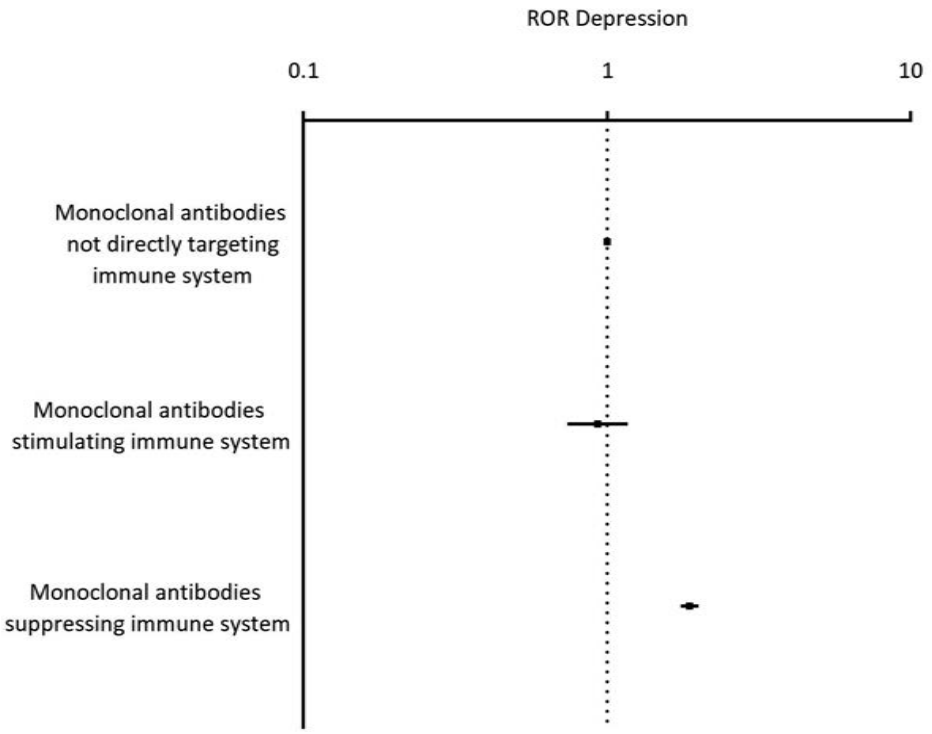


Figure 2. Reporting odds ratios (RORs) relative to monoclonal antibodies not directly targeting the immune system of depression adverse drug reactions (ADRs) for monoclonal antibodies grouped by their influence on the immune system.

the immune system by their intended indication, the ROR for mAbs used for the treatment of autoimmune diseases was higher compared with mAbs not directly targeting the immune system (3.8, 95 CI 3.1–4.7), whereas for mAbs used for the treatment of cancer, this difference was not seen (ROR 1.1, 95% CI 0.7–1.5).

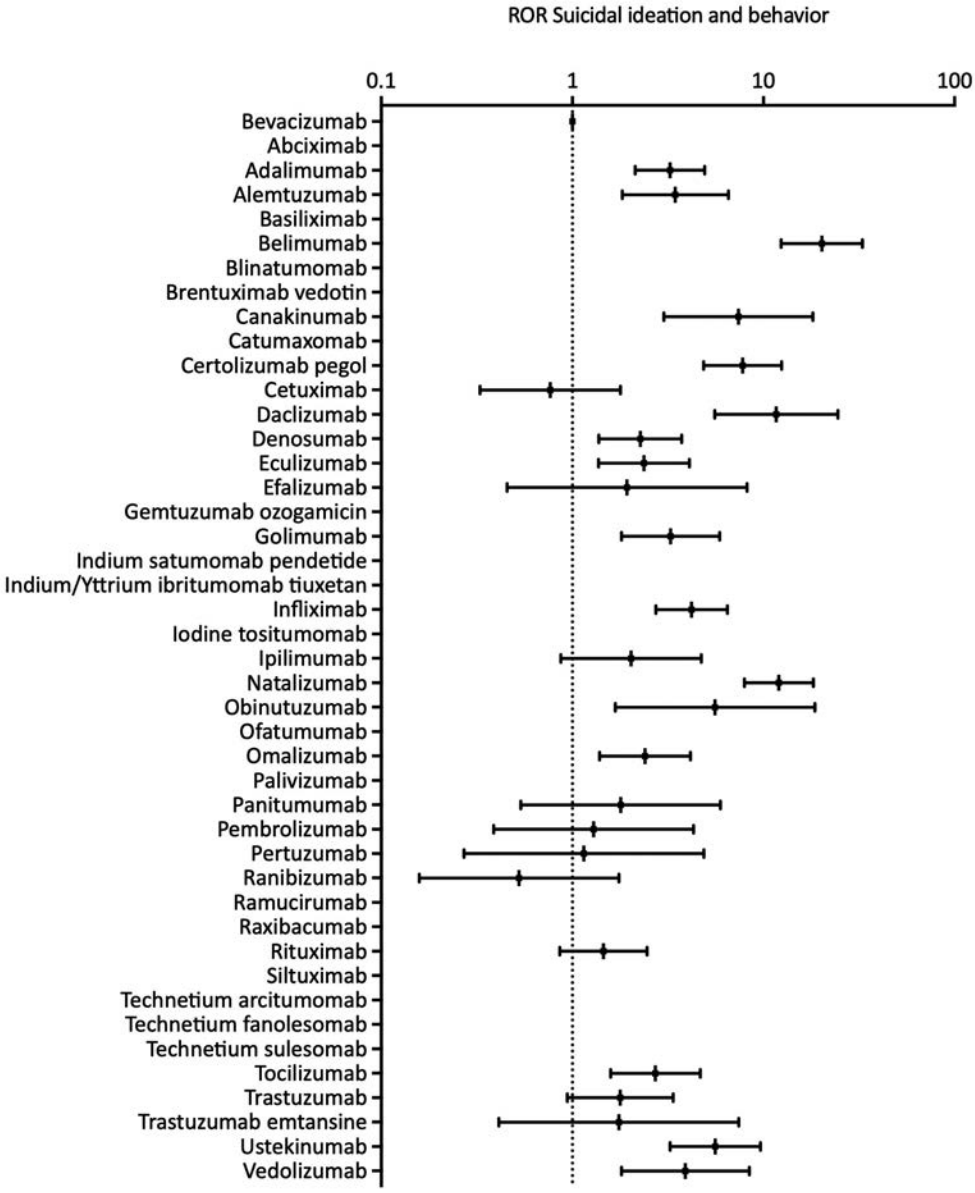


Figure 3. Reporting odds ratios (RORs) relative to bevacizumab of suicidal ideation and behavior adverse drug reactions (ADRs) for the monoclonal antibodies separately.

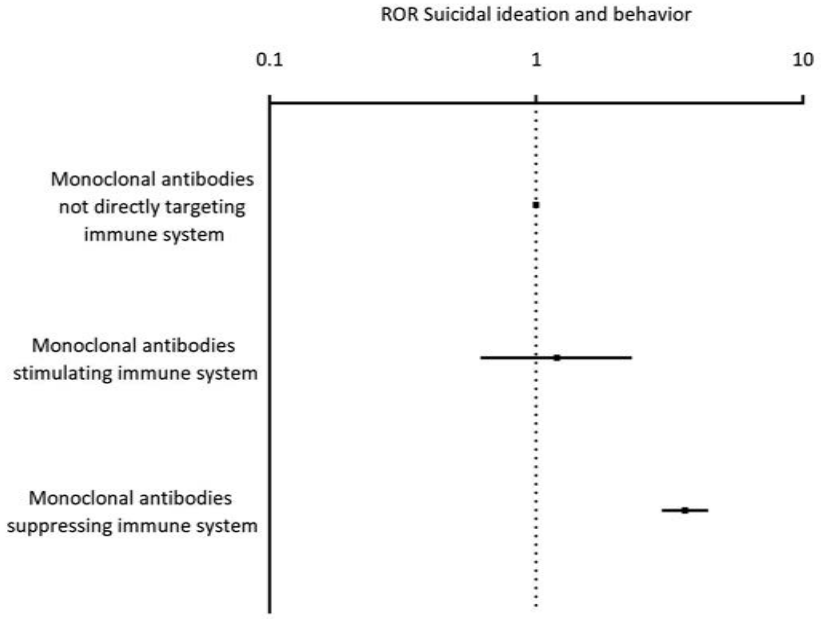


Figure 4. Reporting odds ratios (RORs) relative to monoclonal antibodies not directly targeting the immune system of suicidal ideation and behavior adverse drug reactions (ADRs) for monoclonal antibodies grouped by their influence on the immune system.

DISCUSSION

The current study is, to our knowledge, the first to assess the potential risk of depression and suicidal ideation and behavior for the group of mAbs as a whole. The most relevant finding of this study is that mAb-induced depression and suicidal ideation and behavior seems to be associated with certain specific immune modulating properties.

Depression and suicidal ideation and behavior were most often reported for the mAbs belimumab and natalizumab. When the mAbs were grouped according to their influence on the immune system, we found that depression and suicidal ideation and behavior were more frequently reported for mAbs suppressing the immune system compared with mAbs that do not directly target the immune system. Further stratification by intended indication of the mAbs suppressing the immune system only showed this difference in mAbs used for the treatment of autoimmune diseases. When characterizing the reports, the most marked finding was that the proportion of consumer reports related to depression was approximately twofold higher compared with healthcare professionals.

The discrepancy in the reporting of depression by patients compared with healthcare professionals, as found in this study, is in line with previous studies showing that patients are more likely to report psychiatric ADRs compared with healthcare professionals (28, 29). This may be explained by the nature of the events, as these are experienced by patients to have a direct impact on quality of life.

Belimumab and natalizumab showed the highest reporting of depression and suicidal ideation and behavior relative to bevacizumab. For belimumab, the potential risk of depression was seen in clinical trials and is listed in the product information (30, 31). Furthermore, the pharmaceutical company is currently performing a study to further characterize this potential risk (31). For natalizumab, studies have reported improvement in depression symptoms (32). However, natalizumab has also been reported to cause suicidal ideation and behavior by inducing peripheral cell-mediated inflammation resulting in cytokine secretion (in particular, tumor necrosis factor [TNF]-α) (10). Our study shows a substantial number of reports for natalizumab, indicating that this potential risk should be taken into consideration when treating patients with natalizumab. Besides both agents suppressing the immune system, belimumab and natalizumab do not share further mechanistic commonalities. Due to the limited number of mAbs with more specific mechanistic commonalities (e.g. TNFα inhibitors, IL inhibitors), we were unable to clearly identify patterns in reporting for mAbs with these mechanistic commonalities. As a result, it is not possible to indicate if this potential risk should be monitored for in the post-marketing phase of new mAbs with a specific mode of action.

We also found that depression and suicidal ideation and behavior are more frequently reported for mAbs suppressing the immune system compared with mAbs that do not directly target the immune system. Different meta-analyses showed that inflammatory factors play a role in the pathophysiology of depression and suicidal ideation and behavior (15-20). In general, it is assumed that patients with depression or suicidal ideation and behavior have increased levels of pro-inflammatory factors compared with healthy subjects. However, in meta-analyses, alterations in both pro-inflammatory and anti-inflammatory factors have been linked to these psychiatric disorders. In addition, the influence of psychiatric diseases has been reported for both low and high levels of pro-inflammatory factors (33). The influence exerted by inflammatory factors involves a complicated process and full understanding of their role is lacking. This makes it challenging to fully explain the potential influence of mAbs on these inflammatory factors. In this study, we showed that depression and suicidal ideation and behavior are more frequently reported for mAbs suppressing the immune system compared with mAbs that do not directly target the immune system. This may be explained by the alterations in the levels of inflammatory factors caused by the mAbs.

When the analysis was stratified by the intended indication of the mAb, the difference seen for mAbs suppressing the immune system was only seen in mAbs used for the treatment of autoimmune diseases and not those that suppress the immune system and are used for the treatment of cancer. The potential issue of confounding by indication should therefore be taken into consideration when interpreting these results. As previously described, the estimated prevalence of depression for both patients with cancer and autoimmune diseases exceeds the estimated prevalence in the general population; however, for the main indications of the mAbs suppressing the immune system included in our study, the prevalence of depression seems comparable. In the populations treated with mAbs used for the main autoimmune indications (rheumatoid arthritis, multiple sclerosis, Crohn's disease, psoriasis, and systemic lupus erythematosus), the prevalence of depression is similar and ranges from 20 to 30% (23, 34-37). The main indications for the mAbs used for cancer were

hematological malignancies, and the prevalence of depression is estimated to be approximately 20% in these patients (38).

The strength of this study is that the reports included where retrieved from VigiBase, the largest database containing ICSRs, and were therefore well-suited to obtain insight into reporting patterns, as well as for ADRs with a low reporting frequency, such as suicidal ideation and behavior. Furthermore, the analysis was not only performed for the mAbs separately but aimed to explore the reporting of depression and suicidal ideation and behavior to the mechanism of action of the mAbs.

Limitations

Several limitations are introduced when using data from a spontaneous reporting system. First, we did not have information on the number of patients exposed to the mAbs and the potential history of psychiatric disorders of the patients. In addition, we did not perform a formal causality assessment of the reports. Furthermore, the willingness to report depression and suicidal ideation and behavior may be different between the different disease categories. Therefore, the results of this study should be considered hypothesis-generating, and additional studies are needed to further characterize the potential risk of depression and suicidal ideation and behavior in patients treated with mAbs.

CONCLUSION

Events of depression and suicidal ideation and behavior are reported for mAbs, especially for mAbs that suppress the immune system and that are used for the treatment of autoimmune diseases. This further supports the link between inflammatory factors and the occurrence of depression and suicidal ideation and behavior. The present study provides important knowledge for future research in which, among others, the relation between the different inflammatory factors and the occurrence of depression and suicidal ideation and behavior, indications for use, and population characteristics should be studied to characterize and quantify this potential risk.

ACKNOWLEDGEMENTS

The authors would like to thank the Uppsala Monitoring Centre for providing the data. The information in this paper does not represent the opinion of the World Health Organization, the Uppsala Monitoring Center, or national centers.

REFERENCES

1. Schmidt C. Suicidal thoughts end Amgen's blockbuster aspirations for psoriasis drug. Nat Biotechnol. 2015;33(9):894-5.
2. European Medicines Agency. Kyntheum: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/kyntheum-epar-public-assessment-report_en.pdf. Accessed 6/8/2017.
3. Chen C, Itakura E, Nelson GM, Sheng M, Laurent P, Fenk LA, et al. IL-17 is a neuromodulator of Caenorhabditis elegans sensory responses. Nature. 2017;542(7639):43-8.
4. Silverstein NJ, Huh JR. Interleukin-17: Why the Worms Squirm. Immunity. 2017;46(3):347-9.
5. Choi GB, Yim YS, Wong H, Kim S, Kim H, Kim SV, et al. The maternal interleukin-17a pathway in mice promotes autism-like phenotypes in offspring. Science. 2016;351(6276):933-9.
6. Wallace DJ, Navarra S, Petri MA, Gallacher A, Thomas M, Furie R, et al. Safety profile of belimumab: pooled data from placebo-controlled phase 2 and 3 studies in patients with systemic lupus erythematosus. Lupus. 2013;22(2):144-54.
7. Ellard R, Ahmed A, Shah R, Bewley A. Suicide and depression in a patient with psoriasis receiving adalimumab: the role of the dermatologist. Clin Exp Dermatol. 2014;39(5):624-7.
8. Eshuis EJ, Magnin KM, Stokkers PC, Bemelman WA, Bartelsman J. Suicide attempt in ulcerative colitis patient after 4 months of infliximab therapy--a case report. J Crohns Colitis. 2010;4(5):591-3.
9. Mumoli L, Ciriaco M, Gambardella A, Bombardiere GN, Valentino P, Palleria C, et al. A possible case of natalizumab-dependent suicide attempt: A brief review about drugs and suicide. J Pharmacol Pharmacother. 2013;4(Suppl 1):S90-3.
10. Nagesh O, Bastiampillai T, Fisher L, Mohan T. Cyclical suicidal ideation following natalizumab infusion for multiple sclerosis. Aust N Z J Psychiatry. 2015;49(7):668-9.
11. Roblin X, Oltean P, Heluwaert F, Bonaz B. Panic attack with suicide: an exceptional adverse effect of infliximab. Dig Dis Sci. 2006;51(6):1056.
12. Bagley SC, Sirota M, Chen R, Butte AJ, Altman RB. Constraints on Biological Mechanism from Disease Comorbidity Using Electronic Medical Records and Database of Genetic Variants. PLoS Comput Biol. 2016;12(4):e1004885.
13. Benros ME, Waltoft BL, Nordentoft M, Ostergaard SD, Eaton WW, Krogh J, et al. Autoimmune diseases and severe infections as risk factors for mood disorders: a nationwide study. JAMA Psychiatry. 2013;70(8):812-20.
14. Dantzer R, O'Connor JC, Freund GG, Johnson RW, Kelley KW. From inflammation to sickness and depression: when the immune system subjugates the brain. Nat Rev Neurosci. 2008;9(1):46-56.
15. Köhler CA, Freitas TH, Maes M, de Andrade NQ, Liu CS, Fernandes BS, et al. Peripheral cytokine and chemokine alterations in depression: a meta-analysis of 82 studies. Acta Psychiatr Scand. 2017;135(5):373-87.
16. Liu Y, Ho RC, Mak A. Interleukin (IL)-6, tumour necrosis factor alpha (TNF-α) and soluble interleukin-2 receptors (sIL-2R) are elevated in patients with major depressive disorder: a meta-analysis and meta-regression. J Affect Disord. 2012;139(3):230-9.
17. Eyre HA, Air T, Pradhan A, Johnston J, Lavretsky H, Stuart MJ, et al. A meta-analysis of chemokines in major depression. Prog Neuropsychopharmacol Biol Psychiatry. 2016;68:1-8.
18. Dowlati Y, Herrmann N, Swardfager W, Liu H, Sham L, Reim EK, et al. A meta-analysis of cytokines in major depression. Biol Psychiatry. 2010;67(5):446-57.
19. Hiles SA, Baker AL, de Malmanche T, Attia J. A meta-analysis of differences in IL-6 and IL-10 between people with and without depression: exploring the causes of heterogeneity. Brain Behav Immun. 2012;26(7):1180-8.
20. Haapakoski R, Mathieu J, Ebmeier KP, Alenius H, Kivimäki M. Cumulative meta-analysis of interleukins 6 and 1β, tumour necrosis factor α and C-reactive protein in patients with major depressive disorder. Brain Behav Immun. 2015;49:206-15.

21. European Medicines Agency. Herceptin: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/herceptin-epar-product-information_en.pdf. Accessed 6/8/2017.
22. Pitman A, Suleman S, Hyde N, Hodgkiss A. Depression and anxiety in patients with cancer. Bmj. 2018;361:k1415.
23. Matcham F, Rayner L, Steer S, Hotopf M. The prevalence of depression in rheumatoid arthritis: a systematic review and meta-analysis. Rheumatology (Oxford). 2013;52(12):2136-48.
24. Lindquist M. VigiBase, the WHO Global ICSR Database System: Basic Facts. Drug information journal: DIJ / Drug Information Association. 2008;42(5):409-19.
25. MedDRA Maintenance and Support Services Organization. Medical Dictionary for Regulatory Activities. https://www.meddra.org/. Accessed 29/8/2017.
26. Montastruc JL, Sommet A, Bagheri H, Lapeyre-Mestre M. Benefits and strengths of the disproportionality analysis for identification of adverse drug reactions in a pharmacovigilance database. Br J Clin Pharmacol. 2011;72(6):905-8.
27. van Puijenbroek EP, Bate A, Leufkens HG, Lindquist M, Orre R, Egberts AC. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiol Drug Saf. 2002;11(1):3-10.
28. Banovac M, Candore G, Slattery J, Houÿez F, Haerry D, Genov G, et al. Patient Reporting in the EU: Analysis of EudraVigilance Data. Drug Saf. 2017;40(7):629-45.
29. Avery AJ, Anderson C, Bond CM, Fortnum H, Gifford A, Hannaford PC, et al. Evaluation of patient reporting of adverse drug reactions to the UK 'Yellow Card Scheme': literature review, descriptive and qualitative analyses, and questionnaire surveys. Health Technol Assess. 2011;15(20):1-234, iii-iv.
30. European Medicines Agency. Benlysta: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/benlysta-epar-product-information_en.pdf. Accessed 26/3/2018.
31. European Medicines Agency. Benlysta: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/benlysta-epar-public-assessment-report_en.pdf. Accessed 26/3/2018.
32. Hoepner R, Faissner S, Salmen A, Gold R, Chan A. Efficacy and side effects of natalizumab therapy in patients with multiple sclerosis. J Cent Nerv Syst Dis. 2014;6:41-9.
33. Tonelli LH. Neuroinflammation in suicide: too little may be just as bad as too much. Acta Psychiatr Scand. 2015;131(4):242-3.
34. Dowlatshahi EA, Wakkee M, Arends LR, Nijsten T. The prevalence and odds of depressive symptoms and clinical depression in psoriasis patients: a systematic review and meta-analysis. J Invest Dermatol. 2014;134(6):1542-51.
35. Zhang L, Fu T, Yin R, Zhang Q, Shen B. Prevalence of depression and anxiety in systemic lupus erythematosus: a systematic review and meta-analysis. BMC Psychiatry. 2017;17(1):70.
36. Neuendorf R, Harding A, Stello N, Hanes D, Wahbeh H. Depression and anxiety in patients with Inflammatory Bowel Disease: A systematic review. J Psychosom Res. 2016;87:70-80.
37. Marrie RA, Reingold S, Cohen J, Stuve O, Trojano M, Sorensen PS, et al. The incidence and prevalence of psychiatric disorders in multiple sclerosis: a systematic review. Mult Scler. 2015;21(3):305-17.
38. Linden W, Vodermaier A, Mackenzie R, Greig D. Anxiety and depression after cancer diagnosis: prevalence rates by cancer type, gender, and age. J Affect Disord. 2012;141(2-3):343-51.

SUPPLEMENTARY MATERIAL

Table S1. Grouping of the monoclonal antibodies according to their influence on the immune system.

Influence on the immune system	Monoclonal antibody
Not directly targeting the immune system	Abciximab, bevacizumab, brentuximab vedotin, cetuximab, denosumab, indium satumomab pendetide, palivizumab, panitumumab, pertuzumab, ranibizumab, ramucirumab, raxibacumab, technetium arcitumomab, technetium fanolesomab, technetium sulesomab, trastuzumab, trastuzumab emtansine
Stimulating the immune system	Blinatumomab, catumaxomab, ipilimumab, pembrolizumab
Suppressing the immune system	Adalimumab, alemtuzumab, basiliximab, belimumab, canakinumab, certolizumab pegol, daclizumab, eculizumab, efalizumab, gemtuzumab ozogamicin, golimumab, indium/yttrium ibritumomab tiuxetan, infliximab, iodine tositumomab, natalizumab, obinutuzumab, ofatumumab, omalizumab, rituximab, siltuximab, tocilizumab, ustekinumab, vedolizumab

2.2

THYROID DISORDERS ASSOCIATED WITH PD-1/PD-L1 INHIBITORS: A CLINICAL PRACTICE COHORT STUDY

Lotte A. Minnema, Thijs J. Giezen, Cornelia A.R. Hulsbergen-Veelken,
Mark C.H. de Groot, Karijn P.M. Suijkerbuijk, Toine C.G. Egberts,
Hubert G.M. Leufkens, Helga Gardarsdottir

ABSTRACT

Introduction: Programmed death receptor protein–1/programmed death-ligand 1 (PD-1/PD-L1) inhibitors have shown to be effective in a variety of cancer types, but these drugs can also cause (serious) immune-related adverse events, of which thyroid disorders are the most pronounced.

2.2

Aim: To estimate the incidence, the time to onset of thyroid disorders and time to normalization of the thyroid values in patients treated with PD-1/PD-L1 inhibitors in clinical practice. Furthermore, to assess longitudinal thyroid-level measurement patterns for each patient over time and to explore the patient, disease, and treatment characteristics associated with the occurrence of thyroid disorders.

Methods: Patients treated with PD-1/PD-L1 inhibitors at the University Medical Center Utrecht, the Netherlands, between 2014 and 2019, were included in the cohort. The incidence, time to occurrence of the thyroid disorders and time to normalization of the thyroid values were assessed. Patterns of thyroid hormone levels over time were evaluated using latent profile analysis. A case–control analysis using conditional logistic regression was performed to assess the association between patient, disease, and treatment characteristics and the occurrence of thyroid disorders.

Results: A total of 465 patients were included in this study, of which 13% (n = 58) developed thyroid disorders. Isolated hypothyroidism was observed in 19% (n = 11) of the patients and occurred after a median of 69 days. The remaining 81% (n = 47) of the patients developed hyperthyroidism, which occurred, if isolated, after a median of 55 days (48%, n = 28) and after 21 days for those who subsequently developed hypothyroidism at a median of 48 days later (33%, n = 19). The thyroid levels normalized within a median of 55 days. About 4% (n=14) of the patients experienced a specific pattern with a rapid decline in thyroid-stimulating hormone (TSH) values after initiation of the therapy followed by an increase. This pattern was more prominently observed for female melanoma patients treated with the combination of a PD-1/PD-L1 inhibitor and ipilimumab (4%, n = 14). Female patients and patients with pre-existing thyroid disorders were at increased risk of developing thyroid disorders (odds ratio [OR]: 2.04 [95% confidence interval (CI): 1.14–3.70], and OR: 4.31 [95% CI: 1.47–12.61], respectively).

Conclusion: In patients treated with PD-1/PD-L1 inhibitors, 13% developed thyroid disorders, which occurred more often in female patients and patients with pre-existing thyroid disorders. A specific pattern, with a rapid decline in TSH values followed by an increase, was observed for female melanoma patients treated with a PD-1/PD-L1 inhibitor and ipilimumab. This study also confirmed that thyroid levels should be measured regularly, especially during the first treatment cycles.

INTRODUCTION

Programmed death receptor protein—1/programmed death-ligand 1 (PD-1/PD-L1) inhibitors have emerged as important treatment options in oncology over recent years. The PD-1/PD-L1 inhibitors were studied primarily in patients with advanced melanoma, which is considered to be an immunogenic tumor (1). Currently, the PD-1/PD-L1 inhibitors have shown to be of therapeutic value for the treatment of a large variety of cancer types, such as non-small-cell lung carcinoma, bladder cancer, and classical Hodgkin's lymphoma (2).

The therapeutic effect of the PD-1/PD-L1 inhibitors is achieved through blocking the activity of the PD-1 receptor. As the PD-1 receptor is a negative regulator of T-cell activity, blockade of this receptor potentiates endogenous T-cell responses through which the antineoplastic effect is achieved (3). However, this potentiated immune response can also result in immune-related adverse events. Almost all patients treated with PD-1/PD-L1 inhibitors experience immune-related adverse events, which can range from general adverse events related to the activation of the immune system (e.g., fever, fatigue) to organ-specific immune-related reactions. Some of the most common organ-specific immune-related adverse events are thyroid disorders, including both hypothyroidism and hyperthyroidism. Thyroid disorders occur in approximately 10−-20% of the patients and arise mainly during the first months of treatment (4-10). The underlying mechanism for the occurrence of thyroid disorders is not fully known, with studies presenting conflicting results about the contribution of thyroid-related antibodies (11-15). The management of thyroid disorders differs from that of other immune-related adverse events. Thyroid disorders that require treatment interruptions are rare (5, 16). Moreover, in general, (serious) immune-related adverse events are treated with corticosteroids, whereas the course of thyroid disorders is not thought to be altered by immunosuppressive treatment and they are therefore managed with thyroid hormone replacement therapy only. In addition, while most non-thyroid immune-related adverse events resolve and remain in remission after tapering of immunosuppressants, thyroid disorders are generally irreversible, necessitating chronic thyroid hormone replacement therapy (7). As thyroid disorders are often asymptomatic or present with non-specific symptoms, such as nausea, muscle aches, and tiredness, monitoring of thyroid levels is required during treatment (7, 17, 18).

Previous observational studies identified several risk factors for the occurrence of thyroid disorders. Patel et al. showed that approximately 70% of the patients with abnormalities in thyroid levels prior to initiation of the therapy had an exacerbation of their abnormalities, whereas 35% of the patients with normal thyroid levels prior to the therapy developed abnormalities in thyroid function (19). Sbardella et al. reported that therapy with tyrosine kinase inhibitors is associated with a higher risk of developing hypothyroidism; more than 60% of the patients who developed hypothyroidism were previously treated with a tyrosine kinase inhibitor, compared with 20% of the patients who did not develop hypothyroidism (20). In addition, data from clinical trials has shown a higher predicted incidence of hypothyroidism in patients treated with PD-1 inhibitors compared with patients treated with PD-L1 inhibitors (7.0% vs 3.9%) (5).

Although multiple observational studies have evaluated potential risk factors for the occurrence of thyroid disorders, these studies were limited by the number of patients included. Moreover, these

studies did not fully disentangle the pattern of the thyroid levels over time nor assessed whether populations show different patterns. Therefore, the aims of the current study are to estimate the incidence and time to onset of thyroid disorders and time to normalization of the thyroid values, to describe the pattern of thyroid levels over time, and to explore the patient, disease, and treatment characteristics associated with the occurrence of thyroid disorders in patients treated with PD-1/PD-L1 inhibitors in clinical practice.

METHODS

Study setting

For this observational study, we created a cohort of patients treated with PD-1/PD-L1 inhibitors at the University Medical Center in Utrecht (UMCU), the Netherlands. The UMCU is a 1,042-bed academic teaching hospital in the center of the Netherlands, with annually approximately 28,000 clinical and 15,000 day-care hospitalizations and 334,000 outpatient visits. The UMCU has a specialized cancer center and is appointed as one of the fourteen centers for advanced melanoma treatment in the Netherlands. We used data from the Utrecht Patient Oriented Database (UPOD) for this study. The structure and content of the UPOD have been described in more detail elsewhere (21). In brief, the UPOD is an infrastructure of relational databases comprising data on patient characteristics, hospital-discharge diagnoses, medical procedures, medication orders, and laboratory tests for all patients treated at the UMCU since 2004. Data acquisition and management of the UPOD is in accordance with current regulations concerning privacy and ethics. The UPOD was linked to Cato®, the hospitals' software system used for, among others, prescribing, preparing, and registering the administration of all oncology treatments. For this study, we extracted the patient characteristics (age, sex), medication orders (type of drug, dose, administration date, prescribing date for co-medication), laboratory tests (thyroid hormone values: thyroid-stimulating hormone [TSH] and free thyroxine [FT4], including the date of measurement), and other measurements (weight, length, body mass index [BMI], including the measurement date).

Study population

In the cohort, we included all adult patients (≥18 years) who initiated treatment with the PD-1/PD-L1 inhibitors atezolizumab (Tecentriq®), avelumab (Bavencio®), cemiplimab (Libtayo®), durvalumab (Imfinzi®), nivolumab (Opdivo®), or pembrolizumab (Keytruda®) between July 2014 and December 2019. We included patients only if they had had at least two measurements of the thyroid hormone levels at the UMCU: at least one measurement at the cohort entry date and at least one follow-up measurement within three months following PD-1/PD-L1 treatment initiation. We excluded patients with a history of thyroidectomy and patients with hyperthyroidism or hypothyroidism at the initiation of therapy. We followed patients up until treatment discontinuation or end of the study period (31 December 2019). We considered patients to have discontinued treatment when there was a gap of more than three months without administration of the PD-1/PD-L1 inhibitor. We chose this timeframe because, in the product information of several PD-1/PD-L1 inhibitors, it is described that if, within three months, treatment-related toxicity does not resolve to acceptable grades then

the therapy should be permanently discontinued (17, 22). To calculate the treatment duration, we calculated the time between the first and last administration date with the addition of the cycle duration to the last administration date.

Outcome

The aim of this study was to assess the incidence of thyroid disorders during follow-up within the study population. We measured both the time to occurrence of the thyroid disorders and the time until normalization of the thyroid values. In addition, longitudinal thyroid-level measurement patterns for each patient over time were assessed. We further assessed which patient, disease, and treatment characteristics were associated with the occurrence of thyroid disorders by means of a case–control analysis (see Supplementary information for a graphical depiction of the study design).

Patients were considered to have thyroid disorders when they had at least one measurement of both FT4 and TSH outside the reference range applied at the UMCU during follow-up. Hypothyroidism was defined as an FT4 value of <10 pmol/L and a TSH value of >5.0 mIU/L, and hyperthyroidism was defined as an FT4 value of >22 pmol/L and a TSH value of <0.35 mIU/L. Moreover, patients could develop both hypothyroidism *and* hyperthyroidism during follow-up. According to the clinical guideline, FT4 and TSH values are routinely monitored in patients treated with PD-1/PD-L1 inhibitors. The time to occurrence of the thyroid disorders was defined as the time in days from initiation of PD-1/PD-L1 inhibitor treatment until the first event (event date) during follow-up. The time until the thyroid values normalized was defined as the number of days from the event date until the date when two consecutive measurements of the thyroid levels were again within the reference range. Both the time to occurrence and the time until normalization of the thyroid levels were stratified by the type of adverse event (i.e., isolated hypothyroidism, isolated hyperthyroidism, patients who developed both hypothyroidism and hyperthyroidism). We also assessed whether the patients with thyroid disorders were treated with thyrostatic therapy (carbimazole, thiamazole, propylthiouracil) for hyperthyroidism or thyroid hormone replacement therapy for hypothyroidism and whether this influenced the time until the thyroid values normalized.

To assess the longitudinal thyroid measurements over time, we included patients for whom thyroid levels were available for all of the first four treatment cycles of the PD-1/PD-L1 inhibitor. We used only their TSH values for this analysis, because TSH values are a more sensitive marker for thyroid function than FT4 values. We then logarithmically transformed the TSH values, as these were highly skewed (23).

Patients with thyroid disorders were included as cases in the case–control analysis. Each case was matched to up to five controls from the same cohort using incidence density sampling, thus matching the cases and controls on the basis of the treatment duration (24).

Patient, disease, and treatment characteristics

We included several patient, disease, and treatment characteristics in the case–control analysis to assess whether these were associated with the occurrence of thyroid disorders. The following patient characteristics were included: age (at the cohort entry date), sex, and BMI (measurement

closest to the cohort entry date). Missing values for BMI and/or weight were replaced by the median values in the population calculated based on the sex and indication for which the PD-1/PD-L1 inhibitor was used. Indication for use of the PD-1/PD-L1 inhibitor (lung carcinoma, melanoma, and other: extracted from the description of the diagnosis) was included as a disease characteristic. In addition, pre-existing auto-immune diseases (identified through the recorded diagnosis prior to the cohort entry date) and pre-existing thyroid disorders (identified by prescription of thyroid replacement treatment at any time prior to the cohort entry date) were identified. Finally, the following treatment characteristics were included: the type of PD-1/PD-L1 inhibitor (PD-1 inhibitor [cemiplimab, nivolumab, pembrolizumab] vs PD-L1 inhibitor [atezolizumab, avelumab, durvalumab]), dose (at initiation of the PD-1/PD-L1 inhibitor: below or above the median calculated dose per kilogram), combination therapy (initiated on the same date as the PD-1/PD-L1 inhibitor) with CTLA-4 inhibitor (ipilimumab), concomitant high-dose systemic corticosteroid use in the month prior to the event date (prednisone or equivalent, 1–2 mg/kg/day), co-medication known to affect thyroid function for which the prescription covered the event date (severity of at least grade 2: bexarotene, mitotane, amiodarone, alemtuzumab, interferon-alpha, interleukin-2, highly active antiretroviral therapy, sorafenib, sunitinib, and oral estrogen (25)).

Data analysis

To assess the longitudinal thyroid measurements over time, a latent profile analysis was performed. A discrimination was made in the model with the optimal number of classes based on the Akaike information criterion (AIC) values, for which the model with the lowest AIC value was considered to be the model with the best fit. The characteristics of the patients in the different groups were analyzed using descriptive statistics. The latent profile analysis was performed using the package 'tidyLPA' in R statistical software version 3.6.0 (26).

For the case–control analysis, we calculated the odds ratios (ORs) including 95% confidence intervals (CIs) using univariable conditional logistic regression. A multivariable model was created by including all variables at first and applying stepwise backward selection. Variables were subsequently excluded from the model based on the AIC values. As a sensitivity analysis for the case–control analysis, we altered our hypothyroidism case definition to having a TSH value of >10.0 mIU/L, irrespective of the FT4 value, at any time during follow-up. This is the reference value used in the guideline *Management of Toxicities from Immunotherapy* of the European Society for Medical Oncology, which recommends considering thyroid hormone replacement therapy for the patients with TSH values of >10.0 mIU/L (27).

We performed the data analysis using R statistical software version 3.6.0.

Ethics

The Medical Research Ethics Committee of the UMCU confirmed that the Medical Research Involving Human Subjects Act (WMO) does not apply to this study and that therefore an official ethical approval of the study was not required under the WMO. Data was handled according to the European privacy law and local guidelines.

RESULTS

In total, 545 patients initiated treatment with PD-1/PD-L1 inhibitors. Of these, 80 were excluded as no thyroid levels were available (n = 11), they did not have thyroid levels available at time of initiation of the treatment (n = 48), or these were not available during follow-up (n = 21). As a result, we included 465 patients in this study (Table 1). The mean age of the patients was 63 years (standard deviation [SD]: 12 years) and the majority of the patients were male (61%). Most patients were treated with the PD-1 inhibitors nivolumab (n = 254, 55%) and pembrolizumab (n = 172, 37%). As of 31 December 2019, no patients had been treated with the PD-L1 inhibitor avelumab. Patients were treated with the PD-1/PD-L1 inhibitors for a median duration of 128 days (range: 14–1046 days).

Of the 465 included patients, 13% (n = 58) developed thyroid disorders during follow-up. As shown in Figure 1, isolated hyperthyroidism was observed in 48% (n = 28) of the patients who developed thyroid disorders, and 19% (n = 11) of the patients developed isolated hypothyroidism. The remaining third (33%, n = 19) of the patients first developed hyperthyroidism that later evolved into hypothyroidism. The median time to occurrence of isolated hyperthyroidism was 55 days (range: 8–294 days) and of isolated hypothyroidism was 69 days (range: 14–145 days). For the patients who developed both, the median time to occurrence of hyperthyroidism was 21 days (range: 13–63 days) and the median time to hypothyroidism was 69 days (range: 42–105 days).

Table 1. Baseline characteristics of the cohort of patients treated with PD-1/PD-L1 inhibitors.

Characteristics	n/mean	%/SD
Total	465	
Age in years, mean (SD)	63	12
Body mass index, mean (SD) (n = 353)	25	5
Gender		
Male	282	61%
Drug		
Nivolumab	254	55%
Pembrolizumab	172	37%
Atezolizumab	21	5%
Durvalumab	16	3%
Cemiplimab	2	0%
PD-1 or PD-L1 inhibitor		
PD-1 inhibitor	428	92%
Combination therapy with CTLA-4 inhibitor		
Yes	39	8%
Diagnosis		
Melanoma	192	41%
Lung cancer	137	30%
Other	111	24%
Unknown	25	5%
Pre-existing thyroid disorder		
Yes	22	5%

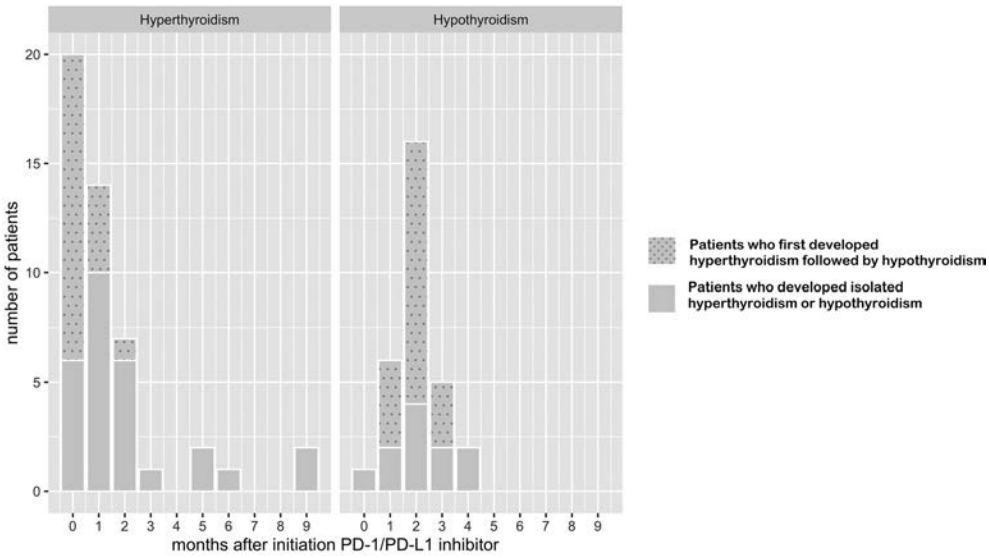


Figure 1. Number of patients who developed hyperthyroidism and hypothyroidism over time.

Of the 30 patients who developed hypothyroidism, 26 were prescribed thyroid hormone replacement therapy. Thyroid hormone replacement therapy was initiated at a median time of two days after diagnosis of the isolated hypothyroidism. Two patients initiated thyroid hormone replacement therapy before both TSH and FT4 were considered to be outside the reference range, which may be explained by the strong increase in TSH values prior to the FT4 levels to be out of the reference range. For the 19 patients who developed hypothyroidism following hyperthyroidism, thyroid hormone replacement therapy was initiated on the day on which the thyroid levels were outside the reference range in 8 patients and ranged from 7 days prior to the development of the hypothyroidism to 21 days after for the other 11 patients. The start of the thyroid hormone replacement therapy in these patients was a median of 46 days (range: 21–84 days) after the diagnosis of hyperthyroidism.

Of the patients who developed hyperthyroidism, one patient was treated with thyrostatic therapy, which was initiated on the day on which the thyroid levels were outside the reference range.

The median time to normalization of the thyroid values was available for 48 patients and was 50 days (range: 14–156 days). For the patients who developed hypothyroidism following initial hyperthyroidism, the median time normalization of the thyroid values was 81 days (range: 23–50 days) from the development of the first event (i.e., hyperthyroidism). For the patients who developed isolated hyperthyroidism or hypothyroidism, the median time to normalization was 42 days. Except from one patient, all patients for whom the time to normalization of the thyroid levels was available were treated with thyroid hormone replacement therapy.

Latent profile analysis

In addition to classifying patients as having hypothyroidism or hyperthyroidism by applying a case definition based on the reference values of the thyroid hormone, patients were also be grouped according to their thyroid values over time. For this, we included the 349 patients for whom thyroid values were available for the first four treatment cycles in the latent profile analysis. The analysis identified four distinct classes of patient exerting different patterns of logarithmic TSH values over time.

As shown in Figure 2 and Table 2, the largest group of patients (64%) had TSH levels within the normal range that did not change during the first four treatment cycles (class 1, stable). The second largest group of patients (36%) showed comparable stable TSH levels, although slightly higher indicating a tendency towards hypothyroidism (class 4, stable-hypo tendency). About 4% of the patients (class 3, hyper-hypo) showed a strong decrease in TSH levels shortly after initiation of the PD-1/PD-L1 inhibitors (i.e., in the second cycle). The TSH values, however, return to normal in the third treatment cycle whereafter the increase in TSH persisted during the last cycle indicating hypothyroidism. The last 8% of the patients (class 2, hyper) showed a decrease in TSH values over time which remained decreasing during each cycle.

As shown in Table 2, the characteristics of patients assigned to class three differed from the characteristics of those in the other three classes. These patients were mostly female melanoma patients treated with nivolumab and a CTLA-4 inhibitor.

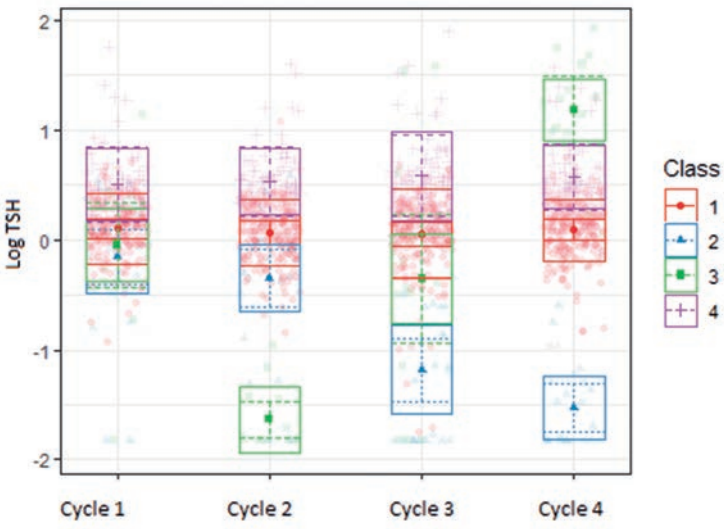


Figure 2. Thyroid-stimulating hormone (TSH) patterns (plotted as the log of the TSH value measured in mIU/L) for the patient classes over the first four treatment cycles.

Table 2. Characteristics of the patients stratified by the assigned class.

Characteristics	Class 1, stable (65%, n = 222)	Class 2, hyper (7%, n = 24)	Class 3, hyper-hypo (4%, n = 14)	Class 4, stable-hypo tendency (26%, n = 89)
Age, mean (SD)	64 (12)	60 (15)	60 (16)	64 (12)
Body mass index, mean (SD)	25 (4)	24 (5)	25 (6)	27 (4)
Gender				
Male	138 (62%)	15 (63%)	5 (36%)	54 (61%)
Diagnosis				
Melanoma	84 (38%)	10 (42%)	8 (57%)	34 (38%)
Lung cancer	75 (34%)	8 (33%)	1 (7%)	19 (21%)
Other	47 (21%)	3 (13%)	5 (36%)	35 (39%)
Unknown	16 (7%)	3 (13%)	0 (0%)	1 (1%)
Drug				
Nivolumab	110 (50%)	11 (46%)	8 (57%)	52 (58%)
Pembrolizumab	94 (42%)	10 (42%)	6 (43%)	31 (35%)
Atezolizumab	6 (3%)	2 (8%)	0 (0%)	5 (6%)
Durvalumab	11 (5%)	0 (0%)	0 (0%)	1 (1%)
Cemiplimab	1 (0.5%)	1 (4%)	0 (0%)	0 (0%)
PD-1 or PD-L1 inhibitor				
PD-1 inhibitor	205 (92%)	22 (92%)	14 (100%)	83 (93%)
Combination therapy with CTLA-4 inhibitor				
Yes	17 (7%)	0 (0%)	4 (29%)	8 (9%)

Case–control analysis

The 58 patients who developed thyroid disorders during follow-up were included as cases in the nested case–control analysis and matched to 290 controls based on the treatment duration (Table 3).

Female patients were demonstrated to be at an increased risk of developing thyroid disorders (Table 3; OR: 2.04 [95% CI: 1.14–3.70]). Patients with pre-existing thyroid disorders had a higher risk of developing thyroid disorders (OR: 4.31 [95% CI: 1.47–12.61]) as compared to patients that did not have pre-existing thyroid disorders. Moreover, patients who were grouped as having a variety of cancer types, including among others Hodgkin's lymphoma, renal cell carcinoma, and urothelial carcinoma, had a higher risk of developing thyroid disorders compared with patients treated with PD-1/PD-L1 inhibitors for lung cancer (OR: 2.27 [95% CI: 1.06–4.87]). Sex, pre-existing thyroid-related adverse events, and the indication for treatment with the PD-1/PD-L1 inhibitors were also the variables that remained in the multivariable model.

A total of six patients were treated with comedication known to affect the thyroid at any time during follow-up. In addition, nine patients were diagnosed with an auto-immune disease prior to the cohort entry date. However, none of these patients developed thyroid disorders. These variables could therefore not be included in the (conditional) case–control analysis.

The sensitivity analysis where the case definition was altered resulted in 16 additional cases. The results of the sensitivity analysis were in line with the main analysis. One additional variable

Table 3. Odds ratios (ORs) from conditional (univariable and multivariable) logistic regression for the association between the covariates and the risk of thyroid disorders. *missing values of weight and body mass index (BMI) were replaced by the median values in the population based on gender and indication.

	Cases	Controls	OR (95% confidence interval); univariable analysis	OR (95% confidence interval); multivariable analysis
Number of patients	58	290		
Age in years, mean (SD)	62 (14)	62 (12)	1.00 (0.97–1.02)	
BMI*, mean (SD)	26 (6)	25 (4)	1.02 (0.96–1.09)	
Sex				
Male	26 (45%)	179 (62%)	Reference	Reference
Female	32 (55%)	111 (38%)	**2.04 (1.14–3.70)**	**1.88 (1.03–3.46)**
Diagnosis				
Lung cancer	12 (21%)	92 (32%)	Reference	Reference
Melanoma	23 (40%)	109 (38%)	1.59 (0.75–3.35)	1.64 (0.77-3.52)
Other	21 (36%)	69 (24%)	**2.27 (1.06–4.87)**	**2.22 (1.03-4.82**
Unknown	2 (3%)	20 (7%)	0.78 (0.16–3.68)	0.92 (0.19-4.41)
PD-1 or PD-L1 inhibitor				
PD-L1 inhibitor	3 (5%)	25 (9%)	Reference	
PD-1 inhibitor	55 (95%)	265 (91%)	1.72 (0.50–5.87)	
Dose**				
Below median dose (kg/mg)	27 (47%)	117 (40%)	Reference	
Median dose or higher (kg/mg)	31 (53%)	173 (60%)	1.15 (0.64–2.05)	
Combination therapy with CTLA-4 inhibitor				
No	50 (86%)	267 (92%)	Reference	
Yes	8 (14%)	23 (8%)	1.83 (0.78–4.28)	
Concomitant high-dose corticosteroid treatment				
No	50 (86%)	252 (87%)	Reference	
Yes	8 (14%)	38 (13%)	1.06 (0.47–2.41)	
Pre-existing thyroid disorder				
No	51 (88%)	280 (97%)	Reference	Reference
Yes	7 (12%)	10 (3%)	**4.31 (1.47–12.61)**	**3.73 (1.25–11.17)**

was demonstrated to be significantly associated with the occurrence of thyroid disorders. Patients treated with a PD-1/PD-L1 inhibitor in combination with a CTLA-4 inhibitor were at increased risk of developing thyroid disorders (OR: 2.60 [95% CI: 1.17–5.81]). However, in the multivariable model this variable did not remain.

DISCUSSION

Within our cohort of patients who initiated treatment with PD-1/PD-L1 inhibitors, 13% developed thyroid disorders during follow-up. The thyroid disorders mainly occurred in the first three months of treatment, with the hyperthyroidism events occurring earlier than the hypothyroidism events, and the thyroid levels generally normalized within two to three months. Most (89%) of the patients who underwent four cycles of treatment presented stable or stable towards hypothyroidism like patterns over time. The remaining two group of patients showed a pattern with an initial strong decline in TSH. The hyper group (7% of patients) showed a pattern with continued decline in TSH values over time, while in the hyper-hypo group (4% of the patients), this decline was followed by a rapid increase in TSH values. Female melanoma patients treated with a PD-1/PD-L1 inhibitor and ipilimumab were overrepresented in the group of patients who showed a strong decline in TSH values followed by a rapid increase. The risk of developing thyroid disorders was increased twofold for female patients as compared to men and fourfold for patients with pre-existing thyroid disorders as compared to patients without pre-existing thyroid disorders. Also, patients categorized as having diagnosis other than melanoma and lung cancer showed to be at increased risk as compared to patients with lung cancer. However, since this is a heterogenous group of patients, the clinical applicability is considered to be limited.

The incidence of thyroid disorders found in our study is in line with other observational studies that applied comparable classifications of thyroid disorders (15, 16, 28). Moreover, in line with other studies, we demonstrated that the thyroid disorders primarily occur during the first months of treatment (29, 30). Although studies have previously described the time to onset and normalization of the thyroid values and have distinguished between isolated hyperthyroidism, isolated hypothyroidism, and patients who developed both, these studies did not formally group patients based on their thyroid levels over time and did not link these to specific patient, disease, and treatment characteristics. With this additional analysis, we showed that in female melanoma patients treated with a PD-1/PD-L1 inhibitor and ipilimumab, the thyroid disorders occur shortly after initiation of the therapy and show a pattern characterized by hyperthyroidism followed by hypothyroidism. For these patients, the initial decrease in TSH was quickly followed by an increase in TSH which could partly explain why thyrostatic drugs were hardly used in this setting. For all patients it remains important to adequately monitor the thyroid function to quickly identify and, if applicable, treat the thyroid disorders. Monitoring is especially relevant in the first months of treatment since the thyroid disorders mainly occurred during this period. This is in line with the European clinical guideline that recommends to assess thyroid function every cycle during the first three months of treatment (27). Latent profile analysis was used to identify the patterns of TSH values and to identify groups of patients exerting similar patterns. Although this analysis did not account for the treatment cycles, which can vary for from two to six weeks, an analysis where we did stratify according to cycle length did not alter our findings.

The strength of this study was that number of patients included in our cohort was greater than that included in other observational studies. However, we did not identify factors that were associated with the risk of developing thyroid disorders additional to the risk factors identified in other

studies. The difficulty of identifying risk factors is not limited to the association with the occurrence of thyroid disorders but is also observed for all immune-related adverse events. A recent study, which used data from the Dutch Melanoma Treatment Registry, assessed the association between patient- and tumor-related factors and severe immune-related adverse events in over 800 patients (31). This study indicated that patients with more advanced disease have a lower risk of developing severe immune-related adverse events; however, no other patient- and tumor-related factors were found to be associated with severe immune-related adverse events (31). Focusing on severe immune-related adverse events, patients with thyroid disorders were underrepresented in that study, explaining the apparently contrasting absence of an association with female gender. Based on the results of other studies, there are potentially several biomarkers for the prediction of immune-related adverse events (32). A study combining data from the US reporting database of adverse events and molecular data demonstrated that several genes related to T-cell activation have the potential to predict the occurrence of immune-related adverse events (33).

For this study, we used data that was collected as part of clinical practice and as such we were dependent on the data available from the electronic hospital care system. The consequent main limitation of this study was that we may have misclassified patients as (not) having pre-existing thyroid disorders, as the data on comedication that is used by the patients is not systematically tracked in the electronic hospital care system. However, given that in 2019 approximately 510,000 people used levothyroxine in the Dutch population, the incidence of 4.7% that we found in our study is higher than the background incidence of levothyroxine use in the Netherlands (34). Moreover, when, for example, a thyroidectomy was performed in another hospital, this may not have been adequately captured in the database. Furthermore, the timing of the measurements of the thyroid values could be of great influence on the outcome of both the time to occurrence of the event and the time to resolution. In our study, the PD-1/PD-L1 inhibitor was administered more than twice without laboratory values available for only 7% (n = 31) of the patients, indicating high compliance with the current clinical guidelines of monitoring thyroid function (27). Given that we retrieved the data from a large academic hospital, 15% (n = 71) of the patients were treated with PD-1/PD-L1 inhibitors as part of a clinical trial. These patients were potentially under stricter control than patients treated in regular clinical practice, although this was not reflected in the number of laboratory values available, which was comparable to the population that was treated as part of clinical practice.

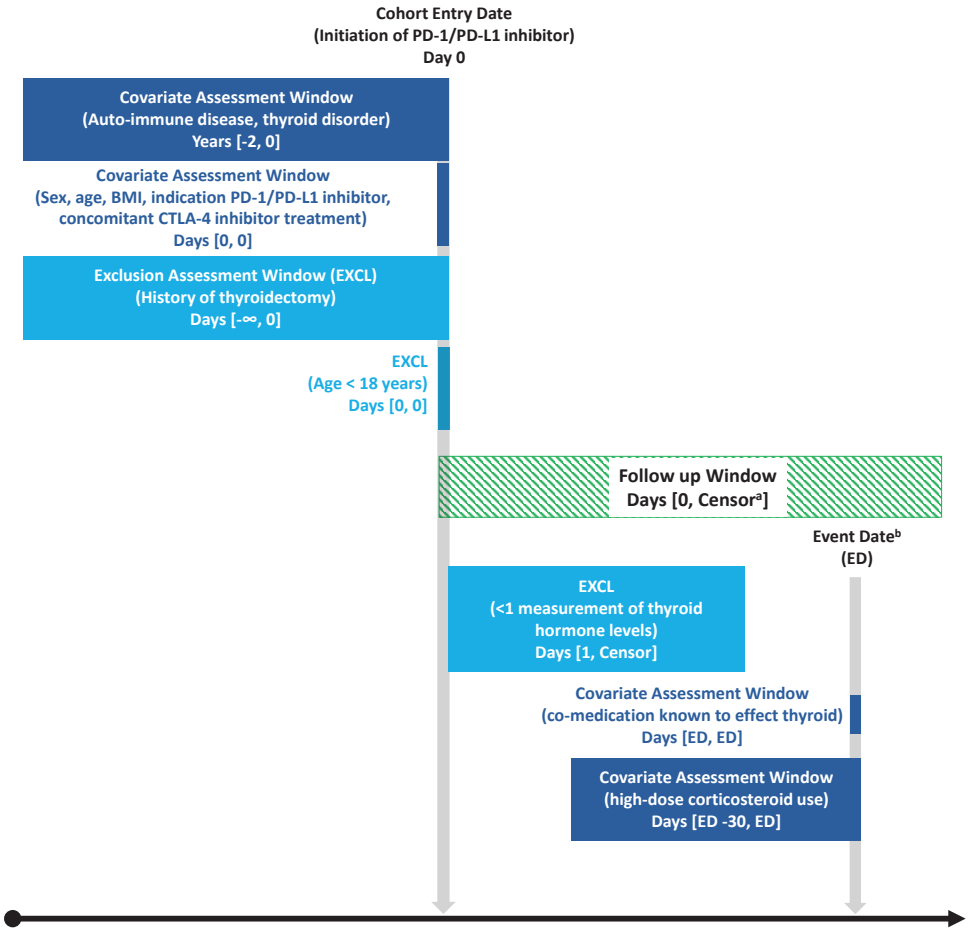
In conclusion, we showed that 13% of the patients treated with PD-1/PD-L1 inhibitors developed thyroid disorders, which more often occurred in women and in patients with pre-existing thyroid disorders. Furthermore, a specific pattern, with a rapid decline in TSH values followed by an increase, was observed for female melanoma patients treated with a PD-1/PD-L1 inhibitor and ipilimumab. This study also confirmed that thyroid levels should be measured regularly, especially during the first treatment cycles.

REFERENCES

1. Sanlorenzo M, Vujic I, Posch C, Dajee A, Yen A, Kim S, et al. Melanoma immunotherapy. Cancer Biol Ther. 2014;15(6):665-74.
2. Gong J, Chehrazi-Raffle A, Reddi S, Salgia R. Development of PD-1 and PD-L1 inhibitors as a form of cancer immunotherapy: a comprehensive review of registration trials and future considerations. J Immunother Cancer. 2018;6(1):8.
3. Alsaab HO, Sau S, Alzhrani R, Tatiparti K, Bhise K, Kashaw SK, et al. PD-1 and PD-L1 Checkpoint Signaling Inhibition for Cancer Immunotherapy: Mechanism, Combinations, and Clinical Outcome. Front Pharmacol. 2017;8:561.
4. Baxi S, Yang A, Gennarelli RL, Khan N, Wang Z, Boyce L, et al. Immune-related adverse events for anti-PD-1 and anti-PD-L1 drugs: systematic review and meta-analysis. BMJ. 2018;360:k793.
5. Barroso-Sousa R, Barry WT, Garrido-Castro AC, Hodi FS, Min L, Krop IE, et al. Incidence of Endocrine Dysfunction Following the Use of Different Immune Checkpoint Inhibitor Regimens: A Systematic Review and Meta-analysis. JAMA Oncol. 2018;4(2):173-82.
6. Sun X, Roudi R, Dai T, Chen S, Fan B, Li H, et al. Immune-related adverse events associated with programmed cell death protein-1 and programmed cell death ligand 1 inhibitors for non-small cell lung cancer: a PRISMA systematic review and meta-analysis. BMC Cancer. 2019;19(1):558.
7. Brahmer JR, Lacchetti C, Schneider BJ, Atkins MB, Brassil KJ, Caterino JM, et al. Management of Immune-Related Adverse Events in Patients Treated With Immune Checkpoint Inhibitor Therapy: American Society of Clinical Oncology Clinical Practice Guideline. J Clin Oncol. 2018;36(17):1714-68.
8. Presotto EM, Rastrelli G, Desideri I, Scotti V, Gunnella S, Pimpinelli N, et al. Endocrine toxicity in cancer patients treated with nivolumab or pembrolizumab: results of a large multicentre study. Journal of Endocrinological Investigation. 2020;43(3):337-45.
9. Jannin A, Penel N, Ladsous M, Vantyghem MC, Do Cao C. Tyrosine kinase inhibitors and immune checkpoint inhibitors-induced thyroid disorders. Crit Rev Oncol Hematol. 2019;141:23-35.
10. Scott ES, Long GV, Guminski A, Clifton-Bligh RJ, Menzies AM, Tsang VH. The spectrum, incidence, kinetics and management of endocrinopathies with immune checkpoint inhibitors for metastatic melanoma. Eur J Endocrinol. 2018;178(2):173-80.
11. Maekura T, Naito M, Tahara M, Ikegami N, Kimura Y, Sonobe S, et al. Predictive Factors of Nivolumab-induced Hypothyroidism in Patients with Non-small Cell Lung Cancer. In Vivo. 2017;31(5):1035-9.
12. Yano S, Ashida K, Nagata H, Ohe K, Wada N, Takeichi Y, et al. Nivolumab-induced thyroid dysfunction lacking antithyroid antibody is frequently evoked in Japanese patients with malignant melanoma. BMC Endocr Disord. 2018;18(1):36.
13. Kobayashi T, Iwama S, Yasuda Y, Okada N, Tsunekawa T, Onoue T, et al. Patients With Antithyroid Antibodies Are Prone To Develop Destructive Thyroiditis by Nivolumab: A Prospective Study. J Endocr Soc. 2018;2(3):241-51.
14. Kimbara S, Fujiwara Y, Iwama S, Ohashi K, Kuchiba A, Arima H, et al. Association of antithyroglobulin antibodies with the development of thyroid dysfunction induced by nivolumab. Cancer Sci. 2018;109(11):3583-90.
15. de Filette J, Jansen Y, Schreuer M, Everaert H, Velkeniers B, Neyns B, et al. Incidence of Thyroid-Related Adverse Events in Melanoma Patients Treated With Pembrolizumab. J Clin Endocrinol Metab. 2016;101(11):4431-9.
16. Kotwal A, Kottschade L, Ryder M. PD-L1 Inhibitor-Induced Thyroiditis Is Associated with Better Overall Survival in Cancer Patients. Thyroid. 2020;30(2):177-84.
17. European Medicines Agency. Keytruda: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/keytruda-epar-product-information_en.pdf. Accessed 1/12/2020.

18. Haanen J, Carbonnel F, Robert C, Kerr KM, Peters S, Larkin J, et al. Management of toxicities from immunotherapy: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2017;28(suppl_4):iv119-iv42.
19. Patel NS, Oury A, Daniels GA, Bazhenova L, Patel SP. Incidence of Thyroid Function Test Abnormalities in Patients Receiving Immune-Checkpoint Inhibitors for Cancer Treatment. Oncologist. 2018;23(10):1236-41.
20. Sbardella E, Tenuta M, Sirgiovanni G, Gianfrilli D, Pozza C, Venneri MA, et al. Thyroid disorders in programmed death 1 inhibitor-treated patients: Is previous therapy with tyrosine kinase inhibitors a predisposing factor? Clinical Endocrinology. 2020;92(3):258-65.
21. ten Berg MJ, Huisman A, van den Bemt PM, Schobben AF, Egberts AC, van Solinge WW. Linking laboratory and medication data: new opportunities for pharmacoepidemiological research. Clin Chem Lab Med. 2007;45(1):13-9.
22. European Medicines Agency. Tecentriq: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/tecentriq-epar-product-information_en.pdf. Accessed 1/12/2020.
23. Sheehan MT. Biochemical Testing of the Thyroid: TSH is the Best and, Oftentimes, Only Test Needed - A Review for Primary Care. Clin Med Res. 2016;14(2):83-92.
24. Vandenbroucke JP, Pearce N. Case-control studies: basic concepts. Int J Epidemiol. 2012;41(5):1480-9.
25. Burch HB. Drug Effects on the Thyroid. N Engl J Med. 2019;381(8):749-61.
26. Rosenberg JM, Beymer PN, Anderson DJ, Van Lissa CJ, Schmidt JA. tidyLPA: An R Package to Easily Carry Out Latent Profile Analysis (LPA) Using Open-Source or Commercial Software. Journal of Open Source Software. 2018;3(30):978.
27. Haanen J, Carbonnel F, Robert C, Kerr KM, Peters S, Larkin J, et al. Management of toxicities from immunotherapy: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2018;29(Suppl 4):iv264-iv6.
28. Basak EA, van der Meer JWM, Hurkmans DP, Schreurs MWJ, Oomen-de Hoop E, van der Veldt AAM, et al. Overt Thyroid Dysfunction and Anti-Thyroid Antibodies Predict Response to Anti-PD-1 Immunotherapy in Cancer Patients. Thyroid. 2020;30(7):966-73.
29. Olsson-Brown A, Lord R, Sacco J, Wagg J, Coles M, Pirmohamed M. Two distinct clinical patterns of checkpoint inhibitor-induced thyroid dysfunction. Endocrine connections. 2020;9(4):318-25.
30. Sakakida T, Ishikawa T, Uchino J, Chihara Y, Komori S, Asai J, et al. Clinical features of immune-related thyroid dysfunction and its association with outcomes in patients with advanced malignancies treated by PD-1 blockade. Oncology letters. 2019;18(2):2140-7.
31. Verheijden RJ, May AM, Blank CU, van der Veldt AAM, Boers-Sonderen MJ, Aarts MJB, et al. Lower risk of severe checkpoint inhibitor toxicity in more advanced disease. ESMO Open. 2020;5(6).
32. Hommes JW, Verheijden RJ, Suijkerbuijk KPM, Hamann D. Biomarkers of Checkpoint Inhibitor Induced Immune-Related Adverse Events-A Comprehensive Review. Front Oncol. 2020;10:585311.
33. Jing Y, Liu J, Ye Y, Pan L, Deng H, Wang Y, et al. Multi-omics prediction of immune-related adverse events during checkpoint immunotherapy. Nat Commun. 2020;11(1):4946.
34. Zorginstituut Nederland. Aantal gebruikers 2015-2019 voor ATC-subgroep H03 : Schildkliermiddelen. https://www.gipdatabank.nl/databank?infotype=g&label=00-totaal&tabel_g_00-totaal=B_01-basis&geg=gebr&spec=&item=H03. Accessed 14/12/2020.
35. Schneeweiss S, Rassen JA, Brown JS, Rothman KJ, Happe L, Arlett P, et al. Graphical Depiction of Longitudinal Study Designs in Health Care Databases. Ann Intern Med. 2019;170(6):398-406.

SUPPLEMENTARY MATERIAL



a. Censored at death, or end of the study period (31-12-2019)
b. Controls risk-set matched on duration of follow up (incidence density sampling)

Figure S1. Visualization of the design of the study (35)

3

DYNAMICS IN POST-MARKETING SAFETY LEARNING

3.1

POST-MARKETING DOSING CHANGES IN THE LABEL OF BIOLOGICALS

Lotte A. Minnema, Thijs J. Giezen, Helga Gardarsdottir,
Toine C.G. Egberts, Hubert G.M. Leufkens, Aukje K. Mantel-Teeuwisse

Br J Clin Pharmacol 2019 Apr;85(4):715-721

ABSTRACT

Aim: The aim of this study was to evaluate post-marketing label changes in dosing information of biologicals.

Methods: Biologicals authorized between 2007 and 2014 by the European Medicines Agency (EMA) were included and followed up from marketing authorization until 31 December 2016 or date of withdrawal of the marketing authorization. The primary outcome of the study was defined as label change in dosing information for the initially approved indication. Incidence of changes, type of change and mean time to change were assessed. As a secondary outcome, label changes in dosing information for extended indications were assessed.

Results: A total of 71 biologicals were included. Dosing information in the label changed for the initial indication during follow-up for eight products (11%). In one of the eight products the change concerned an increase in dose. Also, a change in dosing frequency was identified in three products, for one product a recommendation was added that therapy could be initiated with or without a loading dose, and for one product the minimum dose was removed and a maximum dose was added. For the remaining product the dose was decreased due to safety issues. For 30 products (42%) the indication was extended at least once. No changes in dosing information were observed for the extended indications (n = 59) during follow-up.

Conclusion: This study showed that in 11% of the biologicals, the dosing for the initial indication in the label was changed. In contrast to small molecules, the dose was rarely reduced for safety reasons.

INTRODUCTION

The drug dose aims to optimally balance efficacy, tolerability and safety when treating patients. The drug label, also called the Summary of Product Characteristics in the European Union (EU), informs health care professionals as well as patients about the recommended dose for a given indication. The dose of a (biologic) drug in first-in-man studies is determined based on non-clinical data and subsequently further established in clinical studies. For biologicals, it is different and more difficult to predict their clinical effects from non-clinical data than it is for small molecules because of the complex protein nature of biologicals (1, 2). Specifically, immune reactions such as hypersensitivity reactions and the formation of antidrug antibodies are effects for which prediction by animal models is difficult (3). Also, evidence generation from (non-) clinical trials can be limited by various factors such as the relatively small sample size, the homogeneity of the included population, and the lack of long-term follow-up. Studies conducted after marketing authorization of a new drug, including clinical trials, patient registries and large population-based database studies, aim to provide more information about the efficacy and safety. This post-marketing data can lead to changes in different sections of the label of the product, including the section on dosing information. Previous research showed that the dosing information in the label changed in the post-marketing setting for 21% of new active substances approved by the US Food and Drug Administration (FDA) between 1980 and 1999 (4). In the majority (71%) of the label changes, the dose was reduced, implying that patients may initially be exposed to higher doses than acceptable or needed for the optimal treatment (4). These FDA approval-based findings prompted the European Medicines Agency (EMA) to perform a study on EMA-approved new active substances which showed a comparable frequency in label changes (5). In addition, it was shown that major issues regarding the dose were raised for 10% of the new active substances during the assessment of the marketing application (5).

Dose changes are most often implemented in order to optimize the risk–benefit balance. The ipilimumab example (Box 1) illustrates the difficulties that companies as well as regulators face when finding the dose with the optimal risk–benefit balance for biologicals. Besides increasing the total dose for efficacy-related reasons, the dose can also be increased to prolong the duration of the effect. Due to the pharmacokinetic properties of biologicals, the target can become saturated. In that case the duration of the effect is prolonged (6, 7).

Dose changes can also be introduced as part of the extension of indication. The dosing information for a new therapeutic indication may then differ from the dosing information for the initial indication. For example, rituximab was initially indicated for non-Hodgkin's lymphoma at EU approval in 1998 with a recommended dose of 375 mg m^{-2} body surface area per cycle (8). In 2009, the indication was extended to include another haematological cancer type, chronic lymphocytic leukaemia (8). The recommended dose is 375 mg m^{-2} body surface area in the first cycle followed by 500 mg m^{-2} body surface area in the subsequent cycles. Also, the indication was extended to include a non-oncology indication, rheumatoid arthritis, with a recommended dose of 1000 mg followed by a second 1000 mg 2 weeks later (8).

As described, difficulties are faced in establishing the optimal dose. However, little is known about changes in dosing information for biologicals during the post-marketing phase. Our study

aimed to provide insight into the frequency and nature of post-marketing label changes in dosing for the initial indication of EMA-approved biologicals. Also, changes in the dosing information for the extended indications were assessed.

METHODS

We included biologicals authorized between 1 January 2007 and 31 December 2014 via the centralized procedure of the EMA. According to EMA's definition, biologicals are products produced by or extracted from a biological source (13). We defined biologicals more strictly as recombinant therapeutic (glyco)proteins, thus excluding vaccines, diagnostic proteins, and blood-derived products. Information on the approval circumstance (normal, conditional, under exceptional circumstances) and orphan designation (yes, no) of the biologicals was retrieved from the EMA website. Furthermore, biologicals were classified into the mechanistic classes of antibodies, cytokines, enzymes, growth factors, hormones, interferons, receptors and other/various (14). The product assessment history was retrieved from the EMA website and was used to determine whether a label change in dosing information for the initial indication had occurred and whether the indication was extended. If the assessment history did not provide sufficient information on the occurrence of a label change in dosing, the regulatory assessment report was consulted through the database of the Dutch Medicines Evaluation Board. The biologicals were followed up until 31 December 2016 or until the date of withdrawal if a product was taken off the market.

Incidence of dosing information changes in the drug label, type of dosing information change and time to the dosing information change in the drug label change were assessed. We defined a change in dosing information in the label for the initial indication as an increase or decrease in the dose per dose interval, including increase or decrease in the frequency of administration,

Ipilimumab, a monoclonal antibody activating the immune system by targeting CTLA-4, was approved in the European Union in 2011 for the treatment of advanced melanoma (9). The recommended dose for ipilimumab was 3 mg kg^{-1} every 3 weeks based on the pivotal phase three study. However, there were uncertainties whether the 3 mg kg^{-1} dose induces the maximum pharmacological effect as the pharmacodynamics marker of immune cell activation was increased for the 10 mg kg^{-1} dose compared to the 3 mg kg^{-1} dose. Also, a phase two study had indicated that the 10 mg kg^{-1} dose may be more efficacious though accompanied by an increased number of serious adverse events. As there were multiple differences between those two studies it was not possible to directly compare the results (9). Based on this information it was concluded that it was not fully clear whether 3 mg kg^{-1} is the optimal dose for ipilimumab. Therefore, at approval the regulatory authorities decided that the company should commit to perform a study comparing the efficacy and safety of 3 mg kg^{-1} with 10 mg kg^{-1}. Results of this study became available in 2017 and confirmed that the 10 mg kg^{-1} dose resulted in a significant increase in overall survival compared to the 3 mg kg^{-1} dose, but also in more (serious) adverse events (10-12). The results of this study were included in the label, however, not in the section on dosing information (10).

Box 1. Example difficulties faced in dose tuning.

the dose given per administration, or the duration of the treatment period, and other dose changes (e.g. change in dosing frequency without a change in total dose; 200 mg every 2 weeks changed to 400 mg every 4 weeks). First, the incidence of the occurrence of change in dosing was assessed by dividing the number of changes by the number of biologicals in the cohort. Relative risks, including 95% confidence interval for the occurrence of the first change in dosing for the different determinants, was measured using Cox regression. A Kaplan–Meier analysis was performed to analyse the time to a label change in dosing. If the dosing of a product had changed more than once, only the first change was taken into account for the Kaplan–Meier analysis. The data analysis was performed using SPSS for Windows, version 24.0.

In addition to the changes in dosing information of the initial indication, we determined whether the indication was extended during follow-up. When the indication was extended, the dosing information of the extended indication was compared to the dosing information of the initial indication. The incidence of these differences, type of difference (increase, decrease, other) and time to first extension of indication were assessed. Furthermore, it was assessed whether the dosing information for the extended indications changed during follow-up by comparing the dosing information for the extended indication described in the label at time of the extension of indication to the dosing information in the label for the extended indication at end of follow-up. The labels were obtained from the publicly available community register of medicinal products of the European Commission.

RESULTS

A total of 71 biologicals were included in this study (Supplementary material). Most of the biologicals (n = 64, 90%) were authorized under normal circumstances and did not have an orphan designation (n = 58, 82%). About a third (n = 23, 32%) of the biologicals were hormones, followed by antibodies (n = 22, 31%) and growth factors (n = 10, 14%). Within the follow-up time, a total of five biologicals (pegloticase, rilonacept, filgrastim (n = 2), eptotermin alfa), were withdrawn from the market, all for commercial reasons. Within the median follow-up time of six years (range: 2–10 years), the dosing information in the label for the initial indication was changed for eight products (cumulative incidence: 11%), as shown in Table 1 and Figure 1. The time to the label change ranged from 1 to 7 years after marketing authorization with a median time to a change of 4 years (Figure 2).

For certolizumab and methoxy polyethylene glycolepoetin beta, an alternative dosing regimen with the same total dose was added to the initial dosing regimen. For ranibizumab, the recommended dosing regimen was changed to a less restrictive regimen. For canakinumab and corifollitropin alfa, the dose was increased, whereas for abatacept and romiplostim, the dose was decreased. For romiplostim, the decrease in dose was related to safety. For tocilizumab, the minimum dose was removed and a maximum dose was added. For three products within the cohort, more information about dosing became available after marketing authorization, but the outcomes of these studies did not warrant updates of the recommended dose in the label. We were unable to identify factors related to the label change in dosing information because the sample size was limited.

Table 1. Biologicals whose dosing information was changed in the label for the initial indication.

Biological	Disease category (15)	Description of the label change in dosing information	Time to change (years)
Abatacept	Diseases of the musculoskeletal system and connective tissue	Treatment may be initiated with or without the previously required intravenous loading dose.	6.9
Canakinumab	Diseases of the musculoskeletal system and connective tissue	Increase in the maximum dose from 300 mg or 4 mg/kg every 8 weeks to 600 mg or 8 mg/kg every 8 weeks.	3.3
Certolizumab	Diseases of the musculoskeletal system and c onnective tissue	Addition of an alternative dosing regimen (400 mg every 4 weeks) to the approved dosing regimen (200 mg every 2 weeks) for the treatment of patients with rheumatoid arthritis.	4.2
Corifollitropin alfa	Diseases of the genitourinary system	Increase in dose for patients >36 years and whose weight is between 50 and 60 kg from 100 micrograms to 150 micrograms.	4.8
Methoxy polyethylene glycol-epoetin beta	Diseases of the genitourinary system	Addition of an alternative dosing regimen (0.6 microgram/kg once every two weeks) to the approved dosing regimen (1.2 microgram/ kg once a month) for patients who are not on dialysis and not currently treated with an erythropoiesis stimulating agent.	3.1
Ranibizumab	Diseases of the eye and adnexa	Change in dosing regimen, which is driven by monitoring of the stability of the disease. The initial dosing regimen was based on three initial monthly injections and re-treatment in case of loss of vision.	4.6
Romiplostim	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	Downward revision in cut-off value of thrombocyte count for the recommendation to decrease the dose and to interrupt the treatment.	1.8
Tocilizumab	Diseases of the musculoskeletal system and connective tissue	Removal of the recommendation for a minimum dose (480 mg) and addition of a maximum dose for patients >100 kg (800 mg).	1.4

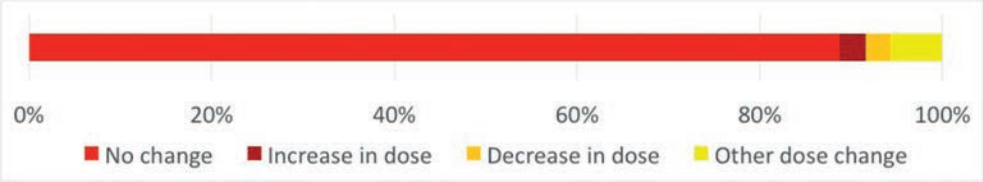


Figure 1. Nature and frequency of label changes in dosing information for the initial indication (n=71).

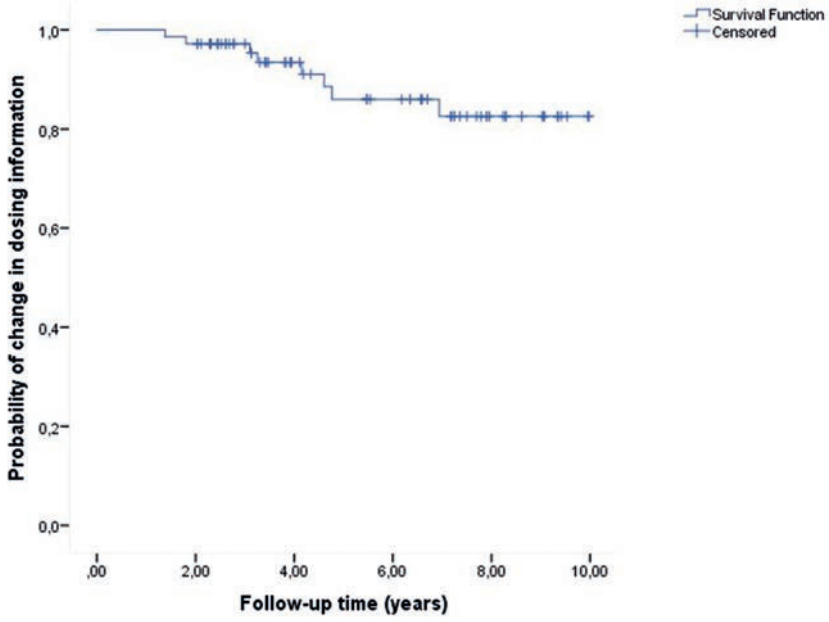


Figure 2. Kaplan Meier curve for the change in the dosing information of the initial indication.

For 30 products (42%), the indication was extended at least once during follow-up with a median time to the first extension of three years (range: 1–7 years). The dose for the extended indication differed from the dose of the initial indication in 15 out of the 30 first extensions of indication (50%), as shown in Figure 3. For 14 products, the indication was extended more than once, resulting in a total of 59 extensions of indication. The dosing for the extended indication differed from the initial dosing in 32 out of these 59 extensions (54%). Furthermore, it was observed that for certolizumab and ranibizumab, the extension of indication was accompanied by a change in dosing information for the initial indication. During follow-up, the dosing information for the extended indications (n = 59) was not changed.

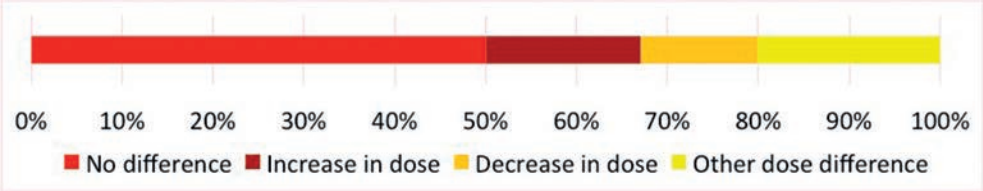


Figure 3. Nature and frequency of differences in dosing information in the label between the initial indication and the first extended indication (n=30).

DISCUSSION

For biologicals, more uncertainties exist about safety at the time of approval than for small molecules (16). However, our study did not show that the dose of biologicals was reduced more often because of safety issues as compared to small molecules. Only one label change included a clear decrease in dose related to safety (romiplostim); the cut-off value warranting a decrease in dose was lowered to minimize the risk of thrombotic complications and was implemented following an international consensus report on the investigation and management of primary immune thrombocytopenia (17). This is in contrast with the previous study on FDA-approved new active substances, which showed that the dose changes occurred more frequently and the dose change was mainly decreased in these products (71%) (4). Four of the changes (abatacept, certolizumab, methoxy polyethylene glycol-epoetin beta, ranibizumab) observed in our study involved (additions of) alternative dosing regimens that reflected a less invasive approach for the patients' convenience (18-21). The remaining three changes (canakinumab, corifollitropin alfa, tocilizumab) were considered efficacy-related changes implemented to optimize the risk–benefit balance. Comparable findings were shown in a study evaluating the rationale of dose selection for FDA-approved biologicals in the pre-approval phase. This study showed that clinical efficacy attributed to the dose finding in 73% of the biologicals, whereas clinical safety attributed in 42% of the biologicals (22).

The extent to which dose changes occur may have been underestimated in our study as in clinical practice dose changes may be introduced based on experience from clinical practice. For example, in rheumatoid arthritis patients treated with TNF-alfa-inhibitors, the dose can effectively be down titrated (23), but down titration is currently not reflected in the label. Moreover, the recommended dose for rituximab in rheumatoid arthritis patients is 1000 mg followed by a second 1000 mg dose 2 weeks later. However, as of today, discussion is still ongoing whether this dose is the optimal dose and in clinical practice patients are often treated with 500 mg instead of 1000 mg (24, 25). More recently, focus in clinical research has also shifted towards tapering of doses for medicines originally recommended for lifelong treatment, e.g. eculizumab, which may have beneficial economic effects (26).

Finally, the dose for the extended indication differed from the dose of the initial indication in half of the first extensions of indication. This indicates that research on dosing continues for extended indications, which may in the end also affect the dosing for the initial indication. In fact, we observed that for two products (certolizumab, ranibizumab), the extension of indication was accompanied by a change in dosing information for the initial indication. In the post-marketing phase, it may

be equally important to emphasize finding the best dose for biologicals from an effectiveness *and* safety perspective rather than from a safety perspective only.

In conclusion, this study showed that in approximately one out of ten EMA-authorized biologicals, the recommended dose changed post-marketing for the initial indication. For the first extended indication, a dose difference between the initial and new indication was observed in one out of two biologicals. The dosing information for the extended indications was not changed during follow-up. In contrast with what previous research has reported for the dose of small molecules, the initial dose of biologicals was almost never reduced for safety reasons.

ACKNOWLEDGEMENTS

The authors wish to acknowledge Dr. Hans C. Ebbers for kindly providing a list of all approved biologicals.

REFERENCES

1. Dixit R, Iciek LA, McKeever K, Ryan PC. Challenges of general safety evaluations of biologics compared to small molecule pharmaceuticals in animal models. Expert Opin Drug Discov. 2010;5(1):79-94.
2. Kizhedath A, Wilkinson S, Glassey J. Applicability of predictive toxicology methods for monoclonal antibody therapeutics: status Quo and scope. Arch Toxicol. 2017;91(4):1595-612.
3. van Meer PJ, Kooijman M, Brinks V, Gispen-de Wied CC, Silva-Lima B, Moors EH, et al. Immunogenicity of mAbs in non-human primates during nonclinical safety assessment. MAbs. 2013;5(5):810-6.
4. Cross J, Lee H, Westelinck A, Nelson J, Grudzinskas C, Peck C. Postmarketing drug dosage changes of 499 FDA-approved new molecular entities, 1980-1999. Pharmacoepidemiol Drug Saf. 2002;11(6):439-46.
5. Ehmann F, Papaluca M, Di Giuseppe F, Pani L, Eskova A, Manolis E, et al. Changes and determination of dosing recommendations for medicinal products recently authorised in the European Union. Expert Opin Pharmacother. 2015;16(6):903-11.
6. Elassaiss-Schaap J, Rossenu S, Lindauer A, Kang SP, de Greef R, Sachs JR, et al. Using Model-Based "Learn and Confirm" to Reveal the Pharmacokinetics-Pharmacodynamics Relationship of Pembrolizumab in the KEYNOTE-001 Trial. CPT Pharmacometrics Syst Pharmacol. 2017;6(1):21-8.
7. Roy A, Mould DR, Wang XF, Tay L, Raymond R, Pfister M. Modeling and simulation of abatacept exposure and interleukin-6 response in support of recommended doses for rheumatoid arthritis. J Clin Pharmacol. 2007;47(11):1408-20.
8. European Medicines Agency. MabThera: EPAR - Product Information. https://www.ema.europa.eu/en/documents/product-information/mabthera-epar-product-information_en.pdf. Accessed 26/1/2019.
9. European Medicines Agency. Yervoy: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/yervoy-epar-public-assessment-report_en.pdf. Accessed 26/1/2019.
10. European Medicines Agency. Yervoy: EPAR - Procedural steps taken and scientific information after authorisation. https://www.ema.europa.eu/en/documents/procedural-steps-after/yervoy-epar-procedural-steps-taken-scientific-information-after-authorisation_en-0.pdf. Accessed 26/1/2019.
11. Ascierto PA, Del Vecchio M, Robert C, Mackiewicz A, Chiarion-Sileni V, Arance A, et al. Ipilimumab 10 mg/kg versus ipilimumab 3 mg/kg in patients with unresectable or metastatic melanoma: a randomised, double-blind, multicentre, phase 3 trial. Lancet Oncol. 2017;18(5):611-22.
12. Menzies AM, Long GV. Optimum dosing of ipilimumab in melanoma: too little, too late? Lancet Oncol. 2017;18(5):558-9.
13. European Commission. Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32001L0083&from=en. Accessed 26/1/2019
14. Giezen TJ, Mantel-Teeuwisse AK, Straus SM, Schellekens H, Leufkens HG, Egberts AC. Safety-related regulatory actions for biologicals approved in the United States and the European Union. Jama. 2008;300(16):1887-96.
15. World Health Organization. International Statistical Classification of Diseases and Related Health Problems 10th Revision. https://icd.who.int/browse10/2015/en. Accessed 26/1/2019.
16. Vermeer NS, Duijnhoven RG, Straus SM, Mantel-Teeuwisse AK, Arlett PR, Egberts AC, et al. Risk management plans as a tool for proactive pharmacovigilance: a cohort study of newly approved drugs in Europe. Clin Pharmacol Ther. 2014;96(6):723-31.
17. Provan D, Stasi R, Newland AC, Blanchette VS, Bolton-Maggs P, Bussel JB, et al. International consensus report on the investigation and management of primary immune thrombocytopenia. Blood. 2010;115(2):168-86.
18. Furst DE, Shaikh SA, Greenwald M, Bennett B, Davies O, Luijtens K, et al. Two dosing regimens of certolizumab pegol in patients with active rheumatoid arthritis. Arthritis Care Res (Hoboken). 2015;67(2):151-60.

19. Schiff M, Weinblatt ME, Valente R, van der Heijde D, Citera G, Elegbe A, et al. Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: two-year efficacy and safety findings from AMPLE trial. Ann Rheum Dis. 2014;73(1):86-94.

20. Regillo CD, Brown DM, Abraham P, Yue H, Ianchulev T, Schneider S, et al. Randomized, double-masked, sham-controlled trial of ranibizumab for neovascular age-related macular degeneration: PIER Study year 1. Am J Ophthalmol. 2008;145(2):239-48.

21. US National Library of Medicine. A Study of Subcutaneous Mircera Once Monthly in the Treatment of Anemia in Participants With Chronic Kidney Disease Not on Dialysis. https://clinicaltrials.gov/ct2/show/results/NCT00559273. Accessed 26/1/2019.

22. Ogasawara K, Breder CD, Lin DH, Alexander GC. Exposure- and Dose-response Analyses in Dose Selection and Labeling of FDA-approved Biologics. Clin Ther. 2018;40(1):95-102.e2.

23. van Herwaarden N, den Broeder AA, Jacobs W, van der Maas A, Bijlsma JW, van Vollenhoven RF, et al. Down-titration and discontinuation strategies of tumor necrosis factor-blocking agents for rheumatoid arthritis in patients with low disease activity. Cochrane Database Syst Rev. 2014(9):Cd010455.

24. Chatzidionysiou K, Lie E, Nasonov E, Lukina G, Hetland ML, Tarp U, et al. Effectiveness of two different doses of rituximab for the treatment of rheumatoid arthritis in an international cohort: data from the CERERRA collaboration. Arthritis Res Ther. 2016;18:50.

25. Schoergenhofer C, Schwameis M, Firbas C, Bartko J, Derhaschnig U, Mader RM, et al. Single, very low rituximab doses in healthy volunteers - a pilot and a randomized trial: implications for dosing and biosimilarity testing. Sci Rep. 2018;8(1):124.

26. Volokhina E, Wijnsma K, van der Molen R, Roeleveld N, van der Velden T, Goertz J, et al. Eculizumab Dosing Regimen in Atypical HUS: Possibilities for Individualized Treatment. Clinical Pharmacology & Therapeutics. 2017;102(4):671-8.

SUPPLEMENTARY MATERIAL

Table S1. List of included biologicals (n=71).

Product	INN
Abasaglar	insulin glargine
Abseamed	epoetin alfa
Accofil	filgrastim
Adcetris	brentuximab vedotin
Arzerra	ofatumumab
Bemfola	follitropin alfa
Benlysta	belimumab
Binocrit	epoetin alfa
Biograstim	filgrastim
Biopoin	epoetin theta
Cimzia	certolizumab pegol
Cyramza	ramucirumab
Elaprase	idursulfase
Elonva	corifollitropin alfa
Entyvio	vedolizumab
Eperzan	albiglutide
Epoetin Hexal	epoetin alfa
Eporatio	epoetin theta
Extavia	interferon beta-1b
Eylea	aflibercept
Fertavid	follitropin beta
Filgrastim Hexal	filgrastim
Filgrastim ratiopharm	filgrastim
Gazyvaro	obinutuzumab
Ilaris	canakinumab
Increlex	mecasermin
Insulin Human Winthrop Rapid	insulin human
Jetrea	ocriplasmin
Kadcyla	trastuzumab emtansine
Krystexxa	pegloticase
Lemtrada	alemtuzumab
Lonquex	lipegfilgrastim
Lucentis	ranibizumab
Mircera	methoxy polyethylene glycol-epoetin beta
Nivestim	filgrastim
NovoEight	turoctocog alfa
Nplate	romiplostim
Nulojix	belatacept
Nuwiq	simoctocog alfa
Opgenra	eptotermin alfa
Orencia	abatacept
Perjeta	pertuzumab
Plegridy	peginterferon beta-1a

Table S1. (continued)

Product	INN
Prolia	denosumab
Ratiograstim	filgrastim
Removab	catumaxomab
Retacrit	epoetin zeta
Revestive	teduglutide
Rilonacept Regeneron	rilonacept
Rixubis	nonacog gamma
RoActemra	tocilizumab
Ruconest	conestat alfa
Ryzodeg	insulin degludec / insulin aspart
Silapo	epoetin zeta
Simponi	golimumab
Soliris	eculizumab
Somatropin Biopartners	somatropin
Stelara	ustekinumab
Sylvant	siltuximab
Tevagrastim	filgrastim
Tresiba	insulin degludec
Trulicity	dulaglutide
Vectibix	panitumumab
Victoza	liraglutide
Vimizim	elosulfase alfa
Vpriv	velaglucerase alfa
Xgeva	denosumab
Xultophy	insulin degludec / liraglutide
Yervoy	ipilimumab
Zaltrap	aflibercept
Zarzio	filgrastim

3.2

REGULATORY SAFETY LEARNING DRIVEN BY THE MECHANISM OF ACTION: THE CASE OF TNF-α INHIBITORS

Lotte A. Minnema, Thijs J. Giezen, Jarno Hoekman, Toine C.G. Egberts,
Hubert G.M. Leufkens, Helga Gardarsdottir

Clinical Pharmacology & Therapeutics 2020 Dec 5. Online ahead of print

ABSTRACT

The summary of product characteristics (SmPC) is an important information source that includes the adverse drug reactions (ADRs) associated with the drug. Drugs with the same mechanism of action are expected to have a similar ADR profile and thus a substantial overlap of the described ADRs in the SmPC. The objective of this study is to assess this overlap. We extracted all ADRs (excl. hypersensitivity and administration site reactions) that were described in the first and all subsequent versions of the SmPCs of all approved TNF-α inhibitors in the European Union. The Medical Dictionary for Regulatory Activities was used to characterize the ADRs. At the end of follow-up, 293 unique ADRs (at high level term level) were described in the SmPCs of the five TNF-α inhibitors. There was substantial variation in the number of ADRs described in the SmPC among the TNF-α inhibitors. Of the 293 ADRs, 133 (45%) were described in the SmPC of one TNF-α inhibitor and 39 (13%) in the SmPCs of all five TNF-α inhibitors. Serious ADRs and ADRs classified as important risks were described approximately four times more often in a second SmPC than ADRs not classified as such. In conclusion, the ADRs described in the SmPCs of the TNF-α inhibitors differ considerably in number and type. In order to adequately inform prescribers and patients, acquired knowledge of the safety profile of drugs with the same mechanism of action should increasingly be taken into account in the assessment of all drugs within the class.

INTRODUCTION

At the time of regulatory approval, the efficacy and safety of a new drug will have been studied in a population of limited size (on average, approximately 1,700 individuals) during a relatively short period of time (1, 2). In addition, the types of patient included in clinical studies can differ considerably from the types of patients using the drug in daily clinical practice, as these clinical studies often exclude patients with multiple diseases and specific populations, such as pregnant women and the elderly (3). Therefore, infrequently occurring adverse drug reactions (ADRs), ADRs that only occur after a long duration of exposure, and ADRs occurring in special populations are usually detected after approval, when the drug is used in daily clinical practice. Regulatory authorities therefore require companies to further characterize their drug's safety profile when it is used in daily clinical practice. This is achieved through collecting and evaluating ADRs and performing post-authorization safety studies. When additional information becomes available, this can result in different regulatory actions. For example, the benefit–risk balance of the drug can become negative due to new information that is collected in clinical practice. This could result in restricting the indication to the patient group for which the benefit–risk balance remains positive or even in revoking the market authorization. A less far-reaching but frequently applied regulatory action is the incorporation of the newly identified safety data in the product information. Annually, hundreds of such safety-related changes in the product information (summary of product characteristics (SmPC) in the European Union (EU)) are approved by the regulatory authorities (4, 5).

The occurrence of many ADRs is a direct consequence of the mechanism of action of a drug. Such ADRs are likely to also occur in users of another drug with the same mechanism of action. For example, artery dissections and aneurysms were found to be associated with the use of vascular endothelial growth factor inhibitors (VEGF-inhibitors) (6, 7). These ADRs are applicable to all VEGF-inhibitors, as VEGF inhibition impairs the vascular wall integrity. Moreover, tuberculosis infections associated with infliximab use were detected during the post-marketing phase and are considered a class effect of the tumor necrosis factor-α inhibitors (TNF-α inhibitors) (8, 9). TNF-α is a pro-inflammatory cytokine that plays a central role in the immune response against tuberculosis infection (8, 9). Treatment with TNF-α inhibitors can therefore reactivate latent tuberculosis infections.

In order to adequately inform prescribers and patients about the safety profile of the drug, ADRs that are linked to the mechanism of action are expected to be described in the product information of all drugs with the same mechanism of action. Previous studies showed a variability in to what extent safety information is presented as part of the regulatory assessment for drugs within the same class. For example, Stefansdottir *et al.* showed that only 40% of the ADRs that were identified in the product information of two drugs within the same class (based on indication, mechanism of action, and structure of the drug) were described in the product information of both drugs (10). Another study showed that serious safety issues identified prior to the approval of HIV drugs were taken into account in the approval process of other drugs within the same class (11). These studies focused either on serious safety events or on a selection of drugs within the same class and therefore did not reflect the complete picture of the dynamics of incorporating class effects in the product information.

For biological medicinal drugs, the occurrence of ADRs is often related to the mechanism of action. Furthermore, biologicals are of specific interest, as over the last decade the share that biologicals comprise of newly launched active substances worldwide has increased and is expected to increase further (12). Therefore, we performed a case study on TNF-α inhibitors, as these represent an important drug class within the group of biologicals because TNF-α inhibitors are considered to be key treatment options for multiple types of autoimmune disease. Furthermore, the first TNF-α inhibitor was approved in the EU in 1999; therefore, extensive experience has been gained with the safety profile of TNF-α inhibitors, allowing early as well as long-term safety aspects to be studied.

The aims of this study are to assess the overlap in ADRs described in the product information of drugs with the same mechanism of action, i.e. TNF-α inhibitors, during the life-cycle of the product, to assess the lag time from the identification of new ADRs to the description of the same ADR in the product information of another TNF-α inhibitor, and to identify factors associated with the description of such ADRs in the product information of multiple TNF-α inhibitors.

METHODS

Extraction, classification, and selection of the ADRs

In this study, we included all ADRs described in the first and all subsequent versions of SmPC of the TNF-α inhibitors that had been approved by the European Commission as of 31 December 2019. We excluded the SmPCs of biosimilars because, according to the EU legislation, the SmPCs of biosimilars are the same as the SmPC of the reference product (13). Although, in the EU, the product information consists of both the SmPC and the patient information leaflet, we focused on the SmPC because the content of the patient information leaflet is directly based on the SmPC. All versions of the SmPCs were retrieved through the Union Register of medicinal products maintained by the European Commission. During the life-cycle of a drug, new information on the efficacy and safety becomes available, which can result in an update to and therefore a new version of the SmPC. Both companies and regulatory authorities can initiate an update of the SmPC. However, irrespective of who initiates an update, it is always assessed by the regulatory authorities. A new version of the SmPC is available in the Union Register when, following the regulatory authorities' recommendation of approval, a positive decision is issued by the European Commission.

All ADRs were extracted using a text-mining method from all versions of the selected SmPCs. The extraction was limited to the specific section of the SmPC that describes the ADRs that are associated with the drug (SmPC section 4.8: "Undesirable effects"). The text-mining method was validated through multiple sources. First, we compared the extracted ADRs with the ADRs available in the PROTECT ADR database. This database was created by the European Medicines Agency (EMA) and partners as part of an Innovative Medicines Initiative funded project (PROTECT Work Package 3) and includes all ADRs described in section 4.8 of all versions of the SmPC up to 30 June 2017 (for more details, see http://www.imi-protect.eu/adverseDrugReactions.shtml). We manually compared the ADRs within the PROTECT database and our extracted ADRs with the ADRs in the SmPC available in the community register of medicinal products of the European Commission. Secondly, through

the text-mining method, it was determined in which version of the SmPC an ADR was first described, which we checked manually for all ADRs.

The ADRs were classified using the Medical Dictionary for Regulatory Activities (MedDRA®) (14). MedDRA® provides validated standardized terminology, which is, among others, used to describe ADRs in the SmPC. MedDRA® has a hierarchical structure. The ADRs in SmPC section 4.8 are usually described at the preferred term level. Preferred terms are grouped into high-level terms (HLTs), which are one level higher in the structure of MedDRA®. For example, the preferred terms "cutaneous tuberculosis" and "pulmonary tuberculosis" fall within the HLT "tuberculous infections". For this study, we assessed the overlap in ADRs described in the SmPCs of the different products at the HLT level, as we considered that this reflects clinical practice most accurately.

We excluded hypersensitivity reactions and administration site reactions, as these are related to the molecule and/or route of administration. Hypersensitivity reactions and administration site reactions were defined as all ADRs included in the Standard MedDRA® query "hypersensitivity" or within the high-level group term (HLGT) "administration site reactions". These include, for example, anaphylactic reactions, administration-related reactions, and Stevens-Johnson syndrome. For the remaining ADRs, we did not specifically assess whether these were considered to be related to the mechanism of action. However, we assumed that these are applicable to all TNF-α inhibitors.

Overlap in the ADRs described in the SmPCs of the different TNF-α inhibitors

The overlap in the ADRs (at the HLT level) described in the SmPCs of the different TNF-α inhibitors was assessed in three ways.

1. *Overlap at initial approval:* At the time of approval of non-first-in-class TNF-α inhibitors, experience will have been gained with the ADRs associated with previously approved TNF-α inhibitors. We assessed for each TNF-α inhibitor (besides the first-in-class) whether ADRs described in the SmPC of previously authorized TNF-α inhibitors were incorporated in the SmPC at the moment of regulatory approval. For example, for the third approved TNF-α inhibitor, we assessed the number of unique ADRs that overlapped with those described in the SmPCs of the first and second approved TNF-α inhibitors. For this, we took into account the latest version of the SmPC of the first and second TNF-α inhibitor prior to approval of the third TNF-α inhibitor. We then assessed whether these ADRs were described in the SmPC at the time of approval of the third TNF-α inhibitor.
2. *Overlap at the end of follow-up:* Extensive experience has been gained of the ADRs associated with the TNF-α inhibitors, given that these have been used in clinical practice for many years. To assess the overlap in ADRs when the safety profile is considered to be mature, we assessed whether an ADR was described in one, two, three, four, or all five of the last versions of the SmPCs of the TNF-α inhibitors (i.e., at the end of follow-up: 31 December 2019).
3. *Lag time in overlap:* When new ADRs are identified for a TNF-α inhibitor, these are considered to be applicable to other TNF-α inhibitors. However, this process takes time. To estimate this lag time, we assessed the time between the first description of an ADR in the SmPC of any of the TNF-α inhibitors (index date) and the uptake of that ADR in the SmPC of another (i.e., a second) TNF-α inhibitor.

Determinants for overlap in the ADRs

The following determinants were assessed to study the overlap in the ADRs described in the SmPCs of the different TNF-α inhibitors.

» *Nature of the ADR:* The ADRs were characterized at the system organ class level, which is the highest level in the MedDRA® hierarchy. We assessed whether the ADR was included in the system organ class infections and infestations.

» *Seriousness of the ADR*: We classified ADRs as being serious if they were included in the important medical events list of the EMA. This list includes the ADRs that result in death, are life threatening, require hospitalization or prolong existing hospitalization, result in persistent or significant disability, or are birth defects (15).

» *Regulatory importance of the ADR:* We categorized the ADR as regulatory important if these were included as such in the risk management plan (RMP) of any of the TNF-α inhibitors at marketing approval or during follow-up. Safety issues are included as important risks in the RMP by the regulators if these should be further characterized after marketing approval and are likely to have an impact on the benefit–risk balance (16). The RMPs were retrieved from the internal database of the Medicines Evaluation Board.

» *First-in-class:* We assessed whether the ADR was described in the SmPC of the first TNF-α inhibitor to be approved within the class of TNF-α inhibitors.

» *Regulatory monitoring:* in the first three years after regulatory approval, the safety of a drug is more frequently evaluated than after this period. We therefore assessed whether ADRs included in the SmPC of, for example, the first and second TNF-α inhibitor were more frequently included in the SmPC of the third TNF-α inhibitor in the first three years after regulatory approval than in the period thereafter.

Within the EU regulatory system, the application assessment is led by different rapporteurs. For each product, we assessed the influence of having at least one rapporteur in common that is responsible for the regulatory assessment on the overlap in ADRs described in the SmPCs of the TNF-α inhibitors at the end of follow-up. Information on the rapporteurs that were responsible for the regulatory assessment of the different TNF-α inhibitors was retrieved from the European Public Assessment Reports published at the time of approval.

Data analysis

We used descriptive statistics to calculate the overlap in the ADRs (at the HLT level) described in the SmPCs of the different TNF-α inhibitors. To assess the overlap at the end of follow-up, we divided the number of ADRs that were described in one, two, three, four or all five of the SmPCs of the different TNF-α inhibitors by the total number of unique ADRs described in the SmPCs at the end of follow-up (31 December 2019). This analysis was also performed at the HLGT level, which is one level higher in the hierarchical structure of MedDRA®. This sensitivity analysis was performed to assess whether comparable HLTs were described in the SmPCs of the TNF-α inhibitors. To assess the overlap at the time of approval, we calculated the number of unique ADRs described in

the SmPCs of other TNF-α inhibitors prior to the approval of the new product. We then calculated the percentage of these ADRs that were described in the SmPC at the time of approval of the product.

Kaplan–Meier analysis was performed to assess the time from the first description of an ADR in the SmPC of any of the TNF-α inhibitors to the time when this ADR was described in the SmPC of another TNF-α inhibitor. As this time cannot be estimated for products that had not been approved at the time when an ADR was first described in an SmPC, we included only the TNF-α inhibitors that had been approved at the time when the ADR was first described.

Hazard ratios (HRs) with corresponding 95% confidence intervals (CIs) were calculated using univariate Cox regression analysis to assess the association between the determinants "nature of the ADR", "seriousness of the ADR", "regulatory importance of the ADR", and "first in class" and the ADR being described in the SmPC of at least two of the TNF-α inhibitors.

To determine the influence of the more intensive regulatory monitoring in the first three years after approval on the overlap, we divided, for each product, the follow-up period in the period following the first three years after approval and the period more than three years after approval. For each product, we assessed the number of ADRs that was first described in the SmPC of the other approved TNF-α inhibitors in both periods (≤3 years after approval and >3 years after approval). We then assessed for each product whether the ADRs that were first described in the SmPC of other TNF-α inhibitors are described in the SmPC of the product in question. This number was then divided by the total number of ADRs identified in the period within and after three years following approval.

To assess the influence of sharing at least one rapporteur responsible for the regulatory assessment, we grouped the TNF-α inhibitors that have at least one rapporteur in common and determined whether, for these TNF-α inhibitors, the overlap in describing the ADRs in the SmPC is different from that for the TNF-α inhibitors that do not share at least one rapporteur.

We performed the data analysis using R statistical software version 3.6.0.

RESULTS

As of 31 December 2019, a total of five TNF-α inhibitors (excluding biosimilars) had been approved in the EU. The first-in-class TNF-α inhibitor (infliximab) was approved in 1999, followed by etanercept, approved in 2000, and adalimumab, approved in 2003. The last two (certolizumab and golimumab) were approved in 2009. None of the drugs was taken off the market during follow-up.

After initial approval, of the five drugs' SmPCs, the SmPC of infliximab was changed the most often (n = 25) to describe new ADRs (at the HLT level, excluding hypersensitivity and administration site reactions), whereas the SmPC of certolizumab was updated three times during follow-up. As shown in Figure 1, there was substantial variation in the number of ADRs described in the SmPCs. At initial approval, a total of 66 ADRs were described in the infliximab SmPC, 41 in the etanercept SmPC, 90 in the adalimumab SmPC, 134 in the certolizumab SmPC, and 73 in the golimumab SmPC. At the end of follow-up, in the adalimumab SmPC the most ADRs were described (n = 200). In the SmPCs of certolizumab, infliximab, etanercept, and golimumab, respectively, 142, 131, 103, and 80 ADRs were described at the end of follow-up. For adalimumab, 110 ADRs were added to the SmPC

after regulatory approval whereas, for certolizumab, eight ADRs were added to the SmPC after regulatory approval.

Overlap in the ADRs described in the SmPCs of the different TNF-α inhibitors

Overlap at initial approval

Prior to the approval of etanercept (second-in-class), experience had been gained with the ADRs associated with infliximab (first-in-class). At the time of the approval of etanercept, a total of 66 unique ADRs were described in the SmPC of infliximab. Of these 66 ADRs, 21 (32%) were described in the initial SmPC of etanercept (Figure 2). Prior to the approval of adalimumab, a total of 90 unique ADRs were described in the SmPCs of infliximab and etanercept. Of these 90 ADRs, 53 (59%) were described in the initial SmPC of adalimumab. Of the 238 unique ADRs that were described in the SmPCs of infliximab, etanercept, and adalimumab prior to the approval of certolizumab and golimumab, 94 ADRs (39%) were described in the initial SmPC of certolizumab whereas 62 (26%) were described in the initial SmPC of golimumab.

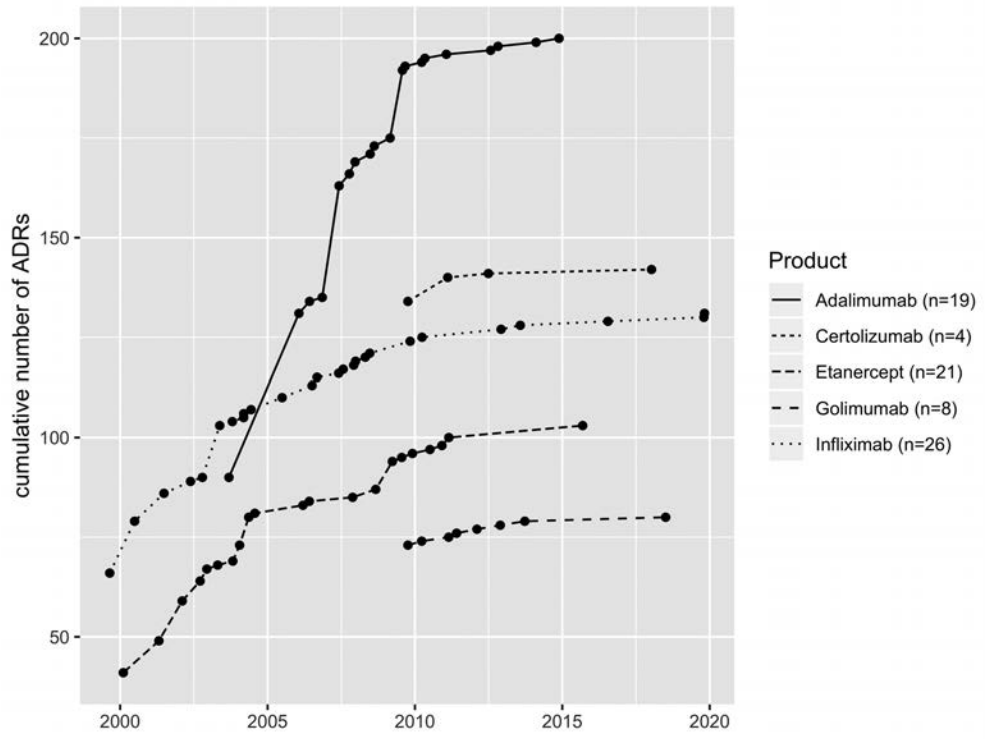


Figure 1. Changes to the SmPCs of the TNF-α inhibitors over time to include new ADRs (at the HLT level, excluding hypersensitivity and administration site reactions); n = number of different versions of the SmPC.

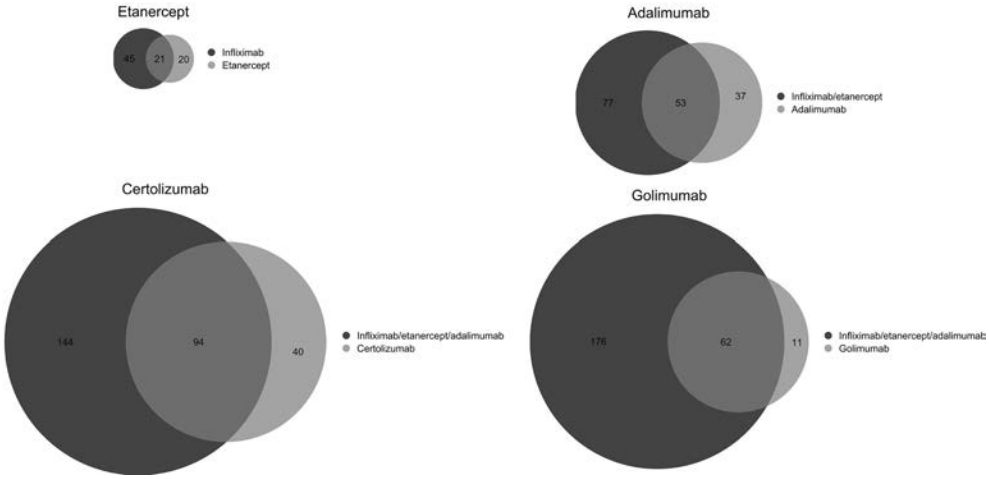


Figure 2. Venn diagrams showing the overlap in the ADRs described in the initial SmPC of etanercept, adalimumab, certolizumab, and golimumab and of the ADRs identified prior to the approval of these drugs.

Overlap at the end of follow-up

At the end of follow-up, a total of 318 different ADRs (at the HLT level) were described in the SmPCs of the TNF-α inhibitors. Of these 318 ADRs, 25 (8%) were classified as hypersensitivity reactions and/or administration site reactions and were therefore disregarded, resulting in a total of 293 ADRs that were included in the analysis. Of these 293 ADRs, 133 (45%) were described in the SmPC of one TNF-α inhibitor, 58 (20%) in the SmPC of two TNF-α inhibitors, 40 (14%) in the SmPC of three TNF-α inhibitors, 23 (8%) in the SmPC of four TNF-α inhibitors, and 39 (13%) in the SmPC of all five TNF-α inhibitors. The 39 ADRs that were included in all SmPCs included tuberculous infections, lower respiratory tract and lung infections, skin melanomas (excluding ocular), and nausea and vomiting symptoms (Table S1).

The sensitivity analysis, performed at one level higher in the hierarchical structure of MedDRA® (HLGT level), showed that, at the end of follow-up, a total of 138 ADRs (HLGTs) were described in the SmPCs. Of these 138 ADRs, 37 (27%) were described in the SmPC of one TNF-α inhibitor, 30 (22%) in the SmPC of two TNF-α inhibitors, 21 (15%) in the SmPC of three TNF-α inhibitors, 20 (14%) in the SmPC of four TNF-α inhibitors, and 30 (22%) in the SmPC of all five TNF-α inhibitors. The 30 ADRs (HLGT level) that were included in all SmPCs included terms such as mycobacterial infectious disorders, general system disorders (not elsewhere classified), and gastrointestinal signs and symptoms (Table S2).

Lag time overlap

One year after the first description of an ADR in the SmPC of any of the TNF-α inhibitors, approximately 7% of these ADRs were described in the SmPC of another TNF-α inhibitor. This percentage increased to approximately 19% after five years. The median lag time between first

description of an ADR in an SmPC to uptake of this ADR in another SmPC was approximately 3 years and ranged from 0 to 15 years (Figure 3).

Determinants for overlap in the ADRs

As shown in Table 1, serious ADRs and ADRs that were classified as important risks by the regulators were described approximately four times more often in the SmPC of at least two TNF-α inhibitors compared with ADRs not classified as such (HR = 4.5, 95% CI: 1.8–10.8; HR = 4.6, 95% CI: 2.0–10.5, respectively). In addition, when the ADR was first described in the SmPC of infliximab (first-in-class), it was described almost three times more often in the SmPC of at least one other TNF-α inhibitors (HR = 2.8, 95% CI: 1.4–5.6) compared to ADRs that were first described in the SmPC of non-first-in-class products. Although ADRs classified as infections and infestations were described more often in a second SmPC compared with other ADRs (HR = 2.1, 95% CI: 1.0–4.5), this difference was not significant.

In the first three years following approval of the individual TNF-α inhibitors, a total of 71 ADRs were first described in the SmPCs of the other TNF-α inhibitors, whereas in the period thereafter a total of 380 ADRs were first described in the SmPCs of other TNF-α inhibitors. Of the 71 ADRs, 31 (44%) were described in the SmPCs of the individual TNF-α inhibitors in the first three years following approval, whereas of the 380 ADRs first described in the SmPCs of the other TNF-α inhibitors >3 years after approval, 25 (7%) were described in the SmPC of the individual TNF-α inhibitors.

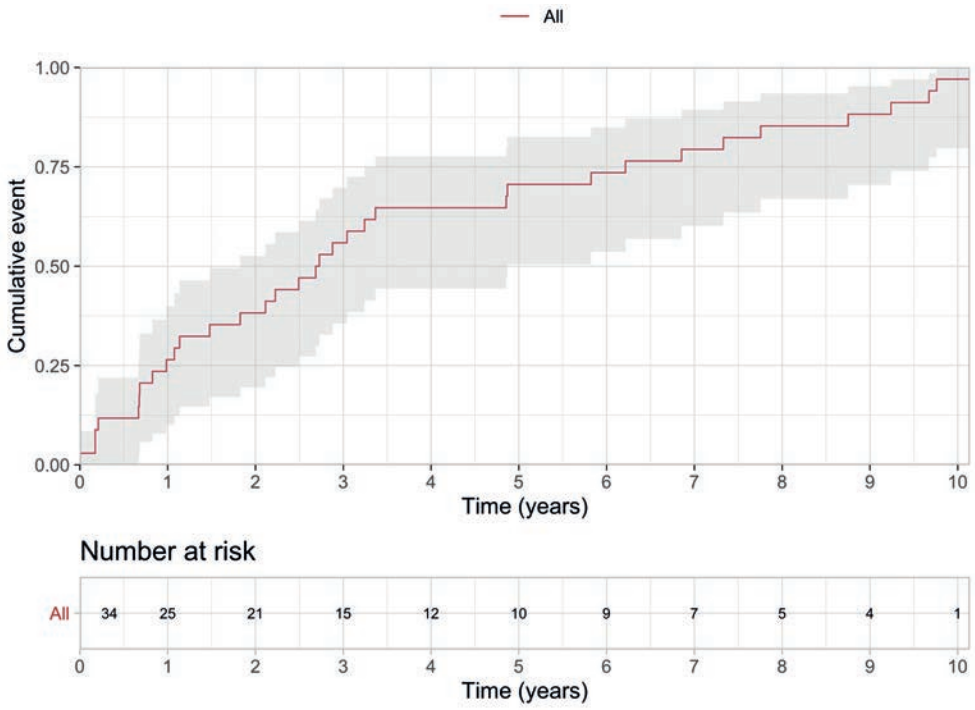


Figure 3. Time from the first description of an ADR in the SmPC of any of the TNF-α inhibitors to the first description of the same ADR in the SmPC of a second TNF-α inhibitor.

Four of the five TNF-α inhibitors shared at least one rapporteur that is responsible for the regulatory assessment. Therefore, the influence of having at least one rapporteur in common on the overlap in ADRs could not be studied.

DISCUSSION

This study showed that the overlap in ADRs (at the HLT level) described in the SmPCs of TNF-α inhibitors is limited; 45% of the ADRs were described in the SmPC of only one TNF-α inhibitor. Moreover, prior knowledge of the ADRs associated with previously approved TNF-α inhibitors is not fully transferred to non-first-in-class TNF-α inhibitors; only 39% of the ADRs that were identified prior to the approval of non-first-in-class TNF-α inhibitors were described in the SmPC at the approval of the non-first-in-class TNF-α inhibitor. If an ADR was described in at least two SmPCs, the median lag time to uptake of the ADR in a second SmPC was approximately 3 years. Specific characteristics of the ADRs ("seriousness", "regulatory importance", and "first-in-class") were shown to be associated with the description of the ADR in at least two SmPCs. As four of the five TNF-α inhibitors shared at least one rapporteur that was responsible for the regulatory assessment, we could not formally study the influence of having a rapporteur responsible for the regulatory assessment in common. However, it can be concluded that although four of the five TNF-α inhibitors shared at least one rapporteur, the overlap of ADRs included in the SmPC of the different TNF-α inhibitors is considered limited. Finally, in the first period after approval, when the regulatory monitoring is more intensive, the percentage of ADRs that was identified and subsequently described in the SmPC was higher than in the period thereafter.

The results of our study are in line with those of other studies that showed that the comparability of the product information of drugs within the same class is limited. Previous studies have shown that the product information differs among regulatory authorities (e.g., the Food and Drug Administration in the US and the EMA in the EU), despite being based on the same information (17, 18). Even between generic drugs that have been proven to be bioequivalent, differences in the product information are present (19). Stefansdottir et al. showed that approximately 40% of ADRs (at the HLGT level) were described in the product information of both the first- and second-in-class drugs (10). Our study showed that approximately 55% of the ADRs at the HLT level and 73% of the ADRs at the HLGT level were described in the product information of at least two TNF-α inhibitors, which is substantially higher. This may be explained by the number of products included, as in our study we included all TNF-α inhibitors while Stefansdottir et al. included the only first- and second-in-class products. We assumed that small differences in the exact mechanism of action did not result in major differences in the safety profile. However, we did not specifically assess whether all ADRs included in our analysis were related to the mechanism of action. For example, ADRs such as headache and nausea are not necessarily related to the mechanism of action but occur generally in patients treated with drugs.

Within the European regulatory system, safety issues are in general assessed on the product level. However, procedures (i.e. signal or referral procedures) are in place in which specific safety issues are evaluated for the group of drugs with the same mechanism of action as a whole. When

the regulatory authorities conclude that an ADR is considered to be a class effect, the ADR should be included in the SmPCs of all drugs involved in the procedure. For example, a signal of lichenoid skin reactions was identified for adalimumab following a scientific publication, and the regulatory authorities concluded that this ADR should be added to adalimumab's SmPC (20). This assessment initiated the evaluation of lichenoid skin reactions as a potential class effect following a literature review, leading to the addition of this ADR in the SmPCs of the other TNF-α inhibitors. In addition, a signal of Kaposi's sarcoma was initially only identified for infliximab based on several reported cases (21). However, based on the data, it was decided that the scope should be extended to all TNF-α inhibitors. These signal and referral procedures are part of the European pharmacovigilance system and can be considered for safety issues identified in the post-marketing setting. However, as illustrated by the results of our study that previous knowledge on the safety profile is not fully taken into account at the time of approval. Therefore, also at the time of approval and as part of extension of indication procedures (potential) class effects should be considered.

The challenge of evaluating ADRs as class effects includes the tradeoff between the level of uncertainty about the causal association and the precaution of adding ADRs to the SmPC that have not (yet) been observed for the specific product. This balance may shift for specific ADRs, as illustrated by the results of our study that serious ADRs and ADRs classified as important risks were significantly more often described in a second SmPC. To facilitate the assessment of class effects, we argue that specific attention should be given to the assessment of the underlying mechanism by constructing adverse outcome pathways. This could follow an integrated approach for which pharmacovigilance data as well as information from clinical and preclinical studies should be taken into account. Also, automated tools are being developed using input from, for example, spontaneous reporting databases and the product information of drugs with the same mechanism of action in order to facilitate the identification of ADRs (22).

Given the nature of the challenge of evaluating ADRs as class effects, this challenge is not limited to the European regulatory system. This is for example illustrated by a study performed using data from the Food and Drug Administration (FDA) showing that major safety issues described in the black box warnings differed among drugs with the same mechanism of action (23). The lag time observed in our study may, however, be different in the US as compared to the EU setting. In the US, ADRs with limited impact on the benefit-risk balance can be submitted to the FDA 30 days prior to distribution of the new product information, whereas updating the SmPC to include new ADRs in the EU typically takes several months (24, 25). Also, the presentation of safety information in the product information differs between the US and EU. For example, in the US product information the incidence of the ADRs observed in the clinical studies is reported for both the experimental and comparator arm, whereas in the EU only the incidence of ADRs in the experimental arm is described. Providing information from both arms, may give further context for health care professionals to the occurrence of the ADRs.

We illustrated the overlap of ADRs described in the SmPCs of drugs with the same mechanism of action on the basis of only one class of drugs. However, considering the earlier-described characteristics of the regulatory system, the results are expected to be applicable to other classes of drug. The European pharmacovigilance system, however, has evolved over time as result of

the implementation of the EU pharmacovigilance legislation that came into force in 2012. The greater part of this study took place prior to this legal change. Nowadays, each product is assigned to a team of (co)rapporteurs from the pharmacovigilance risk assessment committee (PRAC). We expect that this procedure, together with EMA oversight, has resulted in more harmonized SmPCs of recently authorized products compared to those we have studied here. Also, we did not account for differences in indications of the TNF-α inhibitors. Although all TNF-α inhibitors are indicated for rheumatoid arthritis, additional indications differ among the products. The relationship between the number of different indications and the number of ADRs described in the SmPC, however, does not show a consistent pattern. For example, at the initial approval of golimumab and certolizumab, the indication of golimumab was broader than that of certolizumab, whereas approximately 45% more ADRs were described in the certolizumab SmPC compared with the golimumab SmPC (26, 27). Also, the indications of certolizumab and golimumab were extended several times, which did not result in the addition of multiple ADRs, whereas for adalimumab, as part of extension of indication procedures, multiple ADRs were added to the SmPC.

CONCLUSION

Existing as well as new knowledge of ADRs for drugs with the same mechanism of action is not in its entirely described in the safety information of all drugs. Also, when knowledge of ADRs is transferred from one drug to another, this takes considerable time. In order to inform healthcare professionals and patients about the complete picture of the safety profile, knowledge of the safety profile of drugs with the same mechanism of action should increasingly be taken into account for all drugs within the class.

REFERENCES

1. Duijnhoven RG, Straus SM, Raine JM, de Boer A, Hoes AW, De Bruin ML. Number of patients studied prior to approval of new medicines: a database analysis. PLoS Med. 2013;10(3):e1001407.
2. Pregelj L, Verreynne ML, Hine D. Changes in clinical trial length. Nat Rev Drug Discov. 2015;14(5):307-8.
3. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet. 2005;365(9453):82-93.
4. Lester J, Neyarapally GA, Lipowski E, Graham CF, Hall M, Dal Pan G. Evaluation of FDA safety-related drug label changes in 2010. Pharmacoepidemiol Drug Saf. 2013;22(3):302-5.
5. Ebbers HC, Al-Temimi E, Moors EH, Mantel-Teeuwisse AK, Schellekens H, Leufkens HG. Differences between post-authorization adverse drug reactions of biopharmaceuticals and small molecules. BioDrugs. 2013;27(2):167-74.
6. Hatem R, Bebawi E, Schampaert E. Potential Sunitinib-Induced Coronary Artery and Aortic Dissections. Can J Cardiol. 2017;33(6):830.e17-.e18.
7. European Medicines Agency. PRAC recommendation on signals: adopted at the 8-11 July 2019 PRAC meeting. https://www.ema.europa.eu/en/documents/prac-recommendation/prac-recommendations-signals-adopted-8-11-july-2019-prac-meeting_en.pdf. Accessed 3/3/2020.
8. Gardam MA, Keystone EC, Menzies R, Manners S, Skamene E, Long R, et al. Anti-tumour necrosis factor agents and tuberculosis risk: mechanisms of action and clinical management. Lancet Infect Dis. 2003;3(3):148-55.
9. Keane J, Gershon S, Wise RP, Mirabile-Levens E, Kasznica J, Schwieterman WD, et al. Tuberculosis associated with infliximab, a tumor necrosis factor alpha-neutralizing agent. N Engl J Med. 2001;345(15):1098-104.
10. Stefansdottir G, Knol MJ, Arnardottir AH, van der Elst ME, Grobbee DE, Leufkens HG, et al. Safety learning from drugs of the same class: room for improvement. Clin Pharmacol Ther. 2012;91(5):872-80.
11. Arnardottir AH, Haaijer-Ruskamp FM, Straus SM, de Graeff PA, Mol PG. Effect of safety issues with HIV drugs on the approval process of other drugs in the same class: an analysis of European Public Assessment Reports. Drug Saf. 2011;34(11):1101-14.
12. IQVIA Institute for Human Data Science. The Global Use of Medicine in 2019 and Outlook to 2023. https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/the-global-use-of-medicine-in-2019-and-outlook-to-2023.pdf?_=1606729931964. Accessed 3/2/2020.
13. European Medicines Agency. QRD general principles regarding the SmPC information for a generic/hybrid/biosimilar product https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/quality-review-documents-general-principles-regarding-summary-product-characteristics-information/hybrid/biosimilar-product_en.pdf. Accessed 15/3/2020.
14. MedDRA Maintenance and Support Services Organization. Medical Dictionary for Regulatory Activities. https://www.meddra.org/. Accessed 15/3/2020.
15. European Medicines Agency. Inclusion/exclusion criteria for the "Important Medical Events" list. https://www.ema.europa.eu/en/documents/other/eudravigilance-inclusion/exclusion-criteria-important-medical-events-list_en.pdf. Accessed 5/4/2020.
16. European Medicines Agency. Guideline on good pharmacovigilance practices: Module V – Risk management systems (Rev. 2). https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-module-v-risk-management-systems-rev-2_en.pdf. Accessed 5/4/2020.
17. Kesselheim AS, Franklin JM, Avorn J, Duke JD. Speaking the same language? International variations in the safety information accompanying top-selling prescription drugs. BMJ Qual Saf. 2013;22(9):727-34.

18. Eriksson R, Aagaard L, Jensen LJ, Borisova L, Hørlück D, Brunak S, et al. Discrepancies in listed adverse drug reactions in pharmaceutical product information supplied by the regulatory authorities in Denmark and the USA. Pharmacol Res Perspect. 2014;2(3):e00038.

19. Duke J, Friedlin J, Li X. Consistency in the safety labeling of bioequivalent medications. Pharmacoepidemiol Drug Saf. 2013;22(3):294-301.

20. European Medicines Agency. Pharmacovigilance Risk Assessment Committee (PRAC) Minutes of the meeting on 09-12 July 2018. https://www.ema.europa.eu/en/documents/minutes/minutes-prac-meeting-9-12-july-2018_en.pdf. Accessed 18/6/2020.

21. European Medicines Agency. Pharmacovigilance Risk Assessment Committee (PRAC) Minutes of PRAC meeting on 28 – 31 October 2019. https://www.ema.europa.eu/en/documents/minutes/minutes-prac-meeting-28-31-october-2019_en.pdf. Accessed 18/6/2020.

22. Schotland P, Racz R, Jackson DB, Soldatos TG, Levin R, Strauss D, et al. Target adverse event profiles for predictive safety in the post-market setting. Clinical Pharmacology & Therapeutics. 2020 Oct 8. doi: 10.1002/cpt.2074. Online ahead of print

23. Panagiotou OA, Contopoulos-Ioannidis DG, Papanikolaou PN, Ntzani EE, Ioannidis JP. Different black box warning labeling for same-class drugs. J Gen Intern Med. 2011;26(6):603-10.

24. European Commission. Commission Regulation (EU) No 712/2012. https://ec.europa.eu/health//sites/health/files/files/eudralex/vol-1/reg_2012_712/reg_2012_712_en.pdf. Accessed 18/6/2020.

25. Food and Drug Administration. Changes to an Approved NDA or ANDA revision 1. https://www.fda.gov/files/drugs/published/Changes-to-an-Approved-NDA-or-ANDA.pdf. Accessed 18/10/2020.

26. European Medicines Agency. Remicade: EPAR - Scientific Discussion. https://www.ema.europa.eu/en/documents/scientific-discussion/remicade-epar-scientific-discussion_en.pdf. Accessed 15/9/2020.

27. European Medicines Agency. Cimzia: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/cimzia-epar-public-assessment-report_en.pdf. Accessed 15/9/2020.

SUPPLEMENTARY MATERIAL

Table S1. Overview of the high-level terms (grouped by system organ class) described in the SmPC of all TNF-α inhibitors at the end of approval (n = 39).

Blood and lymphatic system disorders
Neutropenias
Anaemias NEC
Marrow depression and hypoplastic anaemias
Leukopenias NEC
Thrombocytopenias

Cardiac disorders
Heart failures NEC

Gastrointestinal disorders
Dyspeptic signs and symptoms
Nausea and vomiting symptoms

General disorders and administration site conditions
Asthenic conditions
Pain and discomfort NEC
Febrile disorders

Hepatobiliary disorders
Hepatocellular damage and hepatitis NEC
Cholecystitis and cholelithiasis

Infections and infestations
Infections NEC
Bacterial infections NEC
Sepsis, bacteraemia, viraemia and fungaemia NEC
Tuberculous infections
Lower respiratory tract and lung infections
Urinary tract infections
Fungal infections NEC
Viral infections NEC
Herpes viral infections
Histoplasma infections
Upper respiratory tract infections
Pneumocystis infections

Investigations
Liver function analyses
Autoimmunity analyses

Musculoskeletal and connective tissue disorders
Lupus erythematosus (incl subtypes)

Table S1. (continued)

Neoplasms benign, malignant and unspecified (incl cysts and polyps)
Skin neoplasms malignant and unspecified (excl melanoma)
Leukaemias NEC
Lymphomas unspecified NEC
Skin melanomas (excl ocular)
Neoplasms malignant site unspecified NEC
Nervous system disorders
Neurological signs and symptoms NEC
Headaches NEC
Skin and subcutaneous tissue disorders
Alopecias
Psoriatic conditions
Vascular disorders
Peripheral embolism and thrombosis
Vascular hypertensive disorders NEC

Table S2. Overview of the high-level group terms (grouped by system organ class) described in the SmPC of all TNF-α inhibitors at the end of approval (n = 30).

Blood and lymphatic system disorders
White blood cell disorders
Anaemias nonhaemolytic and marrow depression
Platelet disorders
Cardiac disorders
Heart failures
Eye disorders
Ocular infections, irritations and inflammations
Gastrointestinal disorders
Gastrointestinal signs and symptoms
Gastrointestinal motility and defaecation conditions
General disorders and administration site conditions
General system disorders NEC
Body temperature conditions
Hepatobiliary disorders
Gallbladder disorders
Hepatic and hepatobiliary disorders
Infections and infestations
Infections - pathogen unspecified

Table S2.(continued)

Bacterial infectious disorders
Fungal infectious disorders
Viral infectious disorders
Mycobacterial infectious disorders
Investigations
Immunology and allergy investigations
Hepatobiliary investigations
Musculoskeletal and connective tissue disorders
Connective tissue disorders (excl congenital)
Neoplasms benign, malignant and unspecified (incl cysts and polyps)
Lymphomas NEC
Miscellaneous and site unspecified neoplasms malignant and unspecified
Skin neoplasms malignant and unspecified
Leukaemias
Nervous system disorders
Neurological disorders NEC
Headaches
Demyelinating disorders
Skin and subcutaneous tissue disorders
Epidermal and dermal conditions
Skin appendage conditions
Vascular disorders
Vascular hypertensive disorders
Embolism and thrombosis

4

SAFETY INFORMATION FROM REGULATORY AND CLINICAL SOURCES

4.1

IDENTIFICATION AND CLASSIFICATION OF THYROID DISORDERS ASSOCIATED WITH PD-1/PD-L1 INHIBITORS IN CLINICAL TRIALS AND OBSERVATIONAL STUDIES: DO THESE DIFFER?

Lotte A. Minnema, Thijs J. Giezen, Toine C.G. Egberts,
Hubert G.M. Leufkens, Helga Gardarsdottir

ABSTRACT

Introduction: Identification and classification of adverse events observed in clinical trials and observational studies is important to facilitate the exchange of safety information. Differences in the identification and classifications of the adverse events between and within clinical trials and observational studies may have implications for the clinical and regulatory interpretation of the safety profile.

Aim: To describe and compare the methods used for the identification and classification of thyroid disorders associated with PD-1/PD-L1 inhibitors in clinical trials and observational studies.

Methods: For the PD-1/PD-L1 inhibitors authorized in the European Union, we identified the pivotal clinical trials that supported the initial approval or extensions of indication. A literature search was performed to identify the observational studies that studied the association between PD-1/L1 inhibitors and the occurrence of thyroid disorders. Information on the methods used to identify patients as having thyroid disorders, subclassifications of thyroid disorders and the classification of severity was retrieved.

Results: We included a total of 38 clinical trials and 28 observational studies. In all of the clinical trials, the method used to identify patients as having thyroid disorders was non-specific as no reference ranges for the thyroid hormones were provided. This method was more specific in the observational studies since the majority of the observational studies (n = 23, 82%) specified the reference ranges for the thyroid hormones. However, these reference ranges differed between the included observational studies. Multiple subclassifications of thyroid disorders were reported with the reporting of subclinical thyroid disorders in the observational studies as main difference in reporting between the clinical trials and observational studies. Moreover, in the clinical trials a specific assessment was performed to subclassify thyroid disorders as being immune-related for which the assessment differed among the different PD-1/PD-L1 inhibitors and, for some, between the clinical trials for one drug. The severity of the adverse events was in all clinical trials classified using the Common Terminology Criteria for Adverse Events, which was also used in 10 (36%) of the observational studies.

Conclusion: Different methods were used to identify and classify thyroid disorders associated with the PD-1/PD-L1 inhibitors. Efforts to improve the comparability of the identification and classification of adverse events should be taken in order to facilitate the regulatory and clinical interpretation of the adverse events.

INTRODUCTION

At the time of marketing approval of a drug, the information on adverse events that occur in patients is mainly based on evidence generated in phase II/III clinical trials. The adverse events experienced by the patients in such trials should be identified by the investigators and subsequently reported by them to the pharmaceutical companies, after which they communicate the adverse events to the regulatory authorities. To facilitate the exchange of safety information, the signs and symptoms of the patients should be identified as adverse events, the adverse events are subclassified and a classification of the severity is performed. In the past, limited regulatory guidance was available leaving it up to the investigators to identify and classify the adverse events. For vaccine trials, for example, it was shown that the identification and classification of adverse events following immunization differed among clinical trials (1). Different cut-off points were used for the identification of fever and the measured body sites differed among the clinical trials (e.g., axillary vs rectal) (1). The comparability was further limited since the time that the patients were observed differed among studies, ranging from 24 – 72 hours to 14 days. Efforts have been made within different disease areas to standardize the identification and classification of adverse events (2-4). Although efforts have been made, the identification and classification of adverse events is still not fully standardized. For example, in clinical trials for Covid-19 vaccines, various cut-off values for temperature were used to identify patients as having fever, ranging from ≥37.8°C to ≥38°C (5-7). The use of variable classifications among different clinical trials may hamper comparisons of the incidences and severity of adverse events among drugs, the performance of meta-analysis, and is of influence on the safety information reported in the regulatory documents and in clinical guidelines. The importance of the uniform classification of adverse events is not limited to the clinical trials. For observational studies, uniform classification of adverse events also facilitates the interpretation of the results of the studies and, if applicable, comparability of results from different studies, which could facilitate the timely detection of emerging safety issues. Moreover, differences in the classification may alter the assessment of whether an adverse event is associated with the drug. For example, it was shown that the association between treatment with antiepileptics and the adverse event suicidality was stronger when suicidality was classified as suicide attempt, suicidal ideation, or deliberate self-harm as compared to when suicidality was classified as completed suicide (8). Besides the differences within clinical trials and observational studies, differences between those may have implications. The safety information described in the product information is namely mainly based on the information collected in the clinical trials. Given that the product information is considered to be important source to guide healthcare professionals and patients on the safe use of the drug, differences in identification and classification of adverse events between clinical trials and clinical practice may have implications. For example, the management of the immune-related adverse events associated with the programmed death receptor protein—1/programmed death-ligand 1 inhibitors (PD-1/PD-L1 inhibitors) is dependent on the classification of the severity of the adverse event and whether the adverse event is considered to be immune-related (9, 10). When these are classified in a different way in the product information and clinical practice, this may have consequences for the adequate identification and management of the adverse events.

The identification and classification of adverse events is extra challenged with the introduction of drugs with new mechanisms of action since the adverse events for these drugs are not yet fully characterized and classifications thereof should be developed. One of these being the introduction of PD-1/PD-L1 inhibitors, which have shown to be effective for a wide variety of cancer types including melanoma and non-small-cell lung cancer (11, 12). The PD-1/PD-L1 inhibitors have a distinct and specific safety profile as compared to other drugs used in oncology namely the occurrence of immune-related adverse events of which thyroid disorders are the most commonly reported (13). Therefore, we aim to describe and compare the identification and classifications of thyroid disorders associated with PD-1/PD-L1 inhibitors in clinical trials and observational studies.

METHODS

Data sources

Different data sources were used to obtain information on the identification and classification of thyroid disorders in clinical trials and observational studies.

Clinical trials

First, all PD-1/PD-L1 inhibitors that were authorized in the European Union as of 30 November 2020 were identified using the website of the European Medicines Agency (EMA). Second, for the included PD-1/PD-L1 inhibitors all pivotal clinical trials supporting initial marketing approval or extensions of indication were retrieved based on the information provided in the European Public Assessment Reports (EPARs) that were available as of 30 November 2020 on EMA's website. The EPAR includes an overview of the assessment procedure and includes all information that supported the initial approval or extension of indication of the drug including the results of the pivotal clinical trials. The corresponding scientific publications of the clinical trials were identified using PubMed and clinicaltrials.gov. The protocols of the clinical trials were retrieved from the supplementary information linked to the scientific publication, clinicaltrials.gov, EPAR, Food and Drug Administration review package, or the company's website.

Observational studies

PubMed was used to identify scientific publications of observational studies assessing the association between PD-1/PD-L1 inhibitor treatment and thyroid disorders (Supplementary information 1). The PubMed search was performed in November 2020. Studies were eligible for inclusion if the study was an observational study (i.e., study making use of clinical practice data), when the identification and classification of thyroid disorders was described and the incidence thereof was or could be calculated. Non-English scientific publications, reviews, and meta-analyses were excluded.

Outcome

The outcomes of this study were the methods used for the identification and classification of thyroid disorders. For this, we extracted information on the assessment used to identify patients as having thyroid disorders, which includes information on the method applied and, if applicable,

the measured laboratory values and reference values of these. Moreover, information was extracted on the reported subclassifications of thyroid disorders (e.g., hypothyroidism and hyperthyroidism) including on whether it was evaluated if other factors could explain the occurrence of the thyroid disorder. Also, information on the classification of the severity of the thyroid disorders was collected. Furthermore, we extracted information on the frequency of performing thyroid function tests, incidence of thyroid disorders, study design, study period, online publication date, number of patients included in the study, PD-1/PD-L1 inhibitor(s) studied, indication of use, dosing schedule, and follow-up time.

We extracted the information for the clinical trials from all information retrieved (e.g., protocols, regulatory information on the assessment procedure). For the observational studies, we extracted the information from the methods section of the full text of the included article.

RESULTS

As of 30 November 2020, six PD-1/PD-L1 inhibitors (atezolizumab, avelumab, cemiplimab, durvalumab, nivolumab, pembrolizumab) were authorized in the European Union. The first PD-1/PD-L1 inhibitors (nivolumab and pembrolizumab) were authorized in 2015 and the last (cemiplimab) in 2019. A total of 38 pivotal clinical trials supported the initial approval or extensions of indications of the six included PD-1/PD-L1 inhibitors (Table 1). The number of clinical trials supporting approval or extension of indication differed among the PD-1/PD-L1 inhibitors ranging from one clinical trial for cemiplimab to 13 clinical trials for pembrolizumab. A total of 28 observational studies were included (Table 1). Also, for the observational studies large variability was observed in the studied drugs: 23 (82%) observational studies included patients treated with nivolumab, 17 (61%) with pembrolizumab, and zero with cemiplimab. One-third of the clinical trials studied the PD-1/PD-L1 inhibitors for the treatment of melanoma, whereas approximately two-third of the observational studies included melanoma patients. The majority of the clinical trials were phase 3 clinical trials (n = 28, 74%) and made use of an active comparator to study the efficacy and safety of the PD-1/PD-L1 inhibitor (n = 25, 66%). All of the observational studies were cohort studies, studying the PD-1/PD-L1 inhibitors without comparator. In the majority of the clinical trials (n = 20, 53%) more than 300 patients were treated with the PD-1/PD-L1 inhibitors. The study population in the observational studies was mostly smaller than the clinical trials with 13 (46%) of the observational studies including less than 100 patients. The treatment duration was generally shorter in the clinical trials as compared to the observational studies, with patients that were treated for less than four months in 13 (34%) of the clinical trials and 2 (7%) of the observational studies. The results of the clinical trials were mostly published in the period between 2014 and 2018, whereas the majority of the results of the observational studies became available between 2017 and 2020.

The method used to identify patients as having thyroid disorder was non-specific in all of the clinical trials, Table 2. Within all of the clinical trials, the Common Terminology Criteria for Adverse Events (CTCAE) version 4(.03) was used to identify patients as having thyroid disorders (Supplementary information 2). Within this dictionary, it is described that thyroid disorders are characterized by excessive levels of thyroid hormone in the body or as a disorder characterized by a decrease in production of thyroid hormone by the thyroid gland (14).

Table 1. Characteristics of the included clinical trials and observational studies. *for the observational studies more than one PD-1/PD-L1 inhibitor and/or indication could be studied.

	Number of clinical trials (%)	Number of observational studies (%)
Total	38	28
PD-1/PD-L1 inhibitor*		
Atezolizumab	8 (21)	3 (11)
Avelumab	2 (5)	1 (4)
Cemiplimab	1 (3)	0 (0)
Durvalumab	2 (5)	1 (4)
Nivolumab	12 (32)	23 (82)
Pembrolizumab	13 (34)	17 (61)
Indication*		
Melanoma	12 (32)	18 (64)
Non-small cell lung cancer	7 (18)	17 (61)
Urothelial carcinoma	6 (16)	5 (18)
Other	13 (34)	10 (36)
Study design		
Phase		
Phase 2	11 (29)	0 (0)
Phase 3	28 (74)	0 (0)
Phase 4 cohort study	0 (0)	28 (100)
Comparator		
Active comparator	25 (66)	0 (0)
Placebo	5 (13)	0 (0)
No comparator	8 (21)	28 (100)
Number of patients treated with PD-1/PD-L1 inhibitor		
<100 patients	1 (3)	13 (46)
100 – 300 patients	17 (45)	13 (46)
>300 patients	20 (53)	2 (7)
Median treatment duration		
<4 months	13 (34)	2 (7)
≥4 months	23 (61)	17 (61)
Online publication date		
2014 – 2016	16 (42)	1 (4)
2017 – 2018	16 (42)	12 (43)
2019 – 2020	6 (16)	15 (54)

For the observational studies, the identification of thyroid disorders included only a description indicating that the specific thyroid levels (only TSH or both TSH and FT4) should be out of the reference range in 5 (18%) of the studies (Supplementary information 3). In 23 (82%) of the observational studies the reference values of TSH and FT4 were specified, Figure 1. The lower limit of the TSH value to classify patients as having thyroid disorders ranged from 0.1 to 0.5 mIU/L and the upper limit of TSH value ranged from 4.2 to 10 mIU/L. Also, the reference values of FT4 differed among the observational studies.

Table 2. Identification and classification of thyroid disorders in clinical trials and observational studies. TSH: thyroid stimulating hormone, ULN: upper limit of normal, LLN: lower limit of normal

	Clinical trials	**Observational studies**
Identification of thyroid disorder		
Non-specific: classified excessive levels of thyroid hormone in the body/decrease in production of thyroid hormone by the thyroid gland	38 (100%)	0 (0%)
Semi specific: classified as increase/decrease TSH value (in combination with increase/ decrease FT4)	0 (0%)	5 (18%)
Specific: cut-off values TSH/FT4 specified	0 (0%)	23 (82%)
Subclassification thyroid disorders	Abnormal thyroid function tests (TSH > ULN: [With at least one FT3/ FT4 test value <LLN] [With all other FT3/FT4 test value >=LLN] TSH < LLN: [With at least one FT3/ FT4 test value >ULN] [With all other FT3/FT4 test value <=ULN]) (Autoimmune) thyroiditis Basedow's disease Hyperthyroidism Hypothyroidism Increased/decreased TSH Immune-related thyroid disorders Thyroxine Free increased/ decreased Triiodothyronine decreased	(Autoimmune/destructive/ biphasic) thyroiditis (Subclinical/overt) hyperthyroidism (Subclinical/overt/new onset) hypothyroidism Isolated thyrotoxicosis Primary overt hyperthyroidism versus thyroiditis-like syndrome Primary thyroid dysfunction Immune-related thyroid dysfunction Recurrent/worsening of pre-existing hypothyroidism Thyroid dysfunction
Classification severity thyroid disorders		
CTCAE dictionary	38 (100%)	10 (36%)

Within the CTCAE dictionary, thyroid disorders are subclassified as hyperthyroidism and hypothyroidism. However, also other subclassifications of thyroid disorders were reported, such as (autoimmune) thyroiditis and increased/decreased TSH. Moreover, in the clinical trials a subclassification of immune-related thyroid disorders including a specific assessment thereof was reported. For this, different assessments were applied for the different PD-1/PD-L1 inhibitors and for atezolizumab, nivolumab, and pembrolizumab the assessments that were applied differed among the clinical trials. In 16 studies, including 11 pembrolizumab trials, it was reported that a selection was made (using the medical dictionary for regulatory activities) to select the adverse events that were

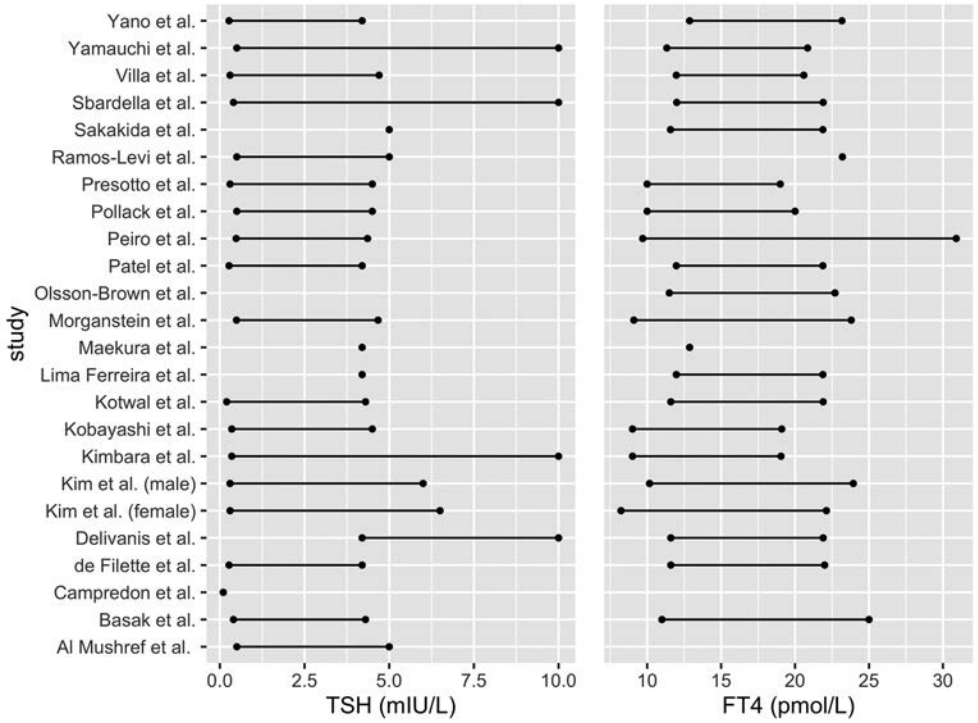


Figure 1. Reference values of thyroid stimulating hormone (TSH) and FT4 applied in the observational studies to identify patients as having thyroid disorders.

(potential) immune-related adverse events. For one nivolumab clinical trial the selected adverse events were all considered to be immune-related, whereas others applied additional criteria to determine the immune-relatedness. In seven studies, the adverse event should have been treated with systemic corticosteroids or, in the case of endocrine adverse events, with endocrine therapy to be subclassified as immune-related adverse event. For avelumab, the most thorough evaluation was performed with two medically-qualified persons subclassifying adverse events as immune-related. In line with the clinical trials, the observational studies reported the subclassifications hyperthyroidism and hypothyroidism. However, the observational studies also made a distinction between subclinical and overt events. For this, thyroid disorders were subclassified as subclinical events when the TSH value was out of the reference range with the FT4 value within the reference range. In the clinical trials, a distinction was made in treatment-related and all-cause adverse events. In two of the observational studies, thyroid disorders were excluded that were considered not to related to the treatment with the PD-1/PD-L1 inhibitor, for which reasons included concurrent neck radiation, use of amiodarone, or that low TSH values were only observed during glucocorticoid therapy.

The frequency of performing thyroid functions tests differed among the studies. In all of the clinical trials thyroid function tests were performed at baseline and in 56% of the clinical trials follow-up measurements were performed before every or every other cycle on treatment. The lowest

frequency of performing thyroid function tests was observed in the pivotal trials supporting initial approval of atezolizumab; thyroid levels were only obtained at baseline and at the end of the study. This frequency was stepped up in the later performed clinical studies of atezolizumab. In a few clinical trials, the frequency of screening was reduced over the course of the clinical trial. The frequency of performing thyroid function tests was specified in 19 of the observational studied and was reported to be performed every or every other cycle.

All of the clinical trials and a third (n = 10, 36%) of the observational studies reported the classification of the severity of the adverse events, for which the classification of the CTCAE dictionary was used. Within the CTCAE dictionary, the severity of the thyroid disorders is classified as grade 1 when the patients are asymptomatic, whereas symptomatic patients are considered to have grade 2 adverse events for which thyroid suppression or replacement therapy is indicated. When patients have severe symptoms, limiting selfcare of activities of daily living, or hospitalization is required adverse events are graded as grade 3. Grade 4 includes events with life-threating consequences or for which urgent intervention is indicated. When the adverse event leads to death, it is considered to be of grade 5.

DISCUSSION

Within this study, we assessed the identification and classification of thyroid disorders associated with PD-1/PD-L1 inhibitors in clinical trials and observational studies and showed that there were both differences between the clinical trials and observational studies as within those. In the clinical trials, the CTCAE dictionary was used to identify patients as having thyroid disorders, which was non-specific and leaves room for interpretation of the investigators. We could therefore not assess whether thyroid disorders were consistently identified in the clinical trials. The CTCAE dictionary could be improved by specifically describing the reference values of the thyroid hormones. The subclassification of immune-related thyroid disorders differed among the PD-1/PD-L1 inhibitors and, for some, also within the clinical trials for one drug. Some trials classified all adverse events that were included on a list of prespecified potential immune-related adverse events (including thyroid disorders) as being immune-related, whereas others used a comparable list as a starting-point for the assessment and, among others, excluded thyroid disorders when they were not treated with endocrine therapy. Since the applied methods for the subclassification of immune-related thyroid disorders differ greatly, the comparability of immune-related adverse events among PD-1/PD-L1 is hampered. Although efforts have been taken to standardize the classification of adverse events by using the CTCAE dictionary, which was updated with immune-related adverse events, such as Guillain-Barre syndrome, myositis, and colitis, the assessment of the subclassification of immune-relatedness is not specified within this dictionary (14). Given that the PD-1/PD-L1 inhibitors are an important class of drugs with a distinct safety profile characterized by immune-related adverse events it would be beneficial to standardize this subclassification. In contrast to the clinical trials, the reference values of the thyroid levels were mostly specified in the observational studies. However, different cut-off values of thyroid levels were used. One contributing factor to facilitate the comparability of the identification of thyroid disorders in observational studies could be

to improve the method to identify thyroid disorders described in clinical practice guidelines. The clinical practice guidelines of the European Society for Medical Oncology and the American Society of Clinical Oncology that describe the management of immune-related adverse events associated with PD-1/PD-L1 inhibitors currently do not describe the reference values (15, 16).

The differences in subclassifications of thyroid disorders could impact the reported incidence and therefore the comparability of the incidences within and between clinical trials and observational studies. Especially, including subclinical events when determining the incidence of thyroid disorders is of influence. For example, Kim et al. and Peiro et al. reported an incidence of thyroid disorders of 32.8% and 23.3% of which approximately half consisted of patients that were classified as having a subclinical event. This difference of including subclinical events to determine the incidence of thyroid disorders is not limited to the observational studies and is also observed in the information provided in the product information of the PD-1/PD-L1 inhibitors, which is based on the clinical trials. For example, the incidence of hypothyroidism described in the product information of atezolizumab is calculated based on reports that include, among others, blood thyroid stimulating hormone abnormal and blood thyroid stimulating hormone increased, and thyroiditis (17). Other PD-1/PD-L1 inhibitors determine the incidence of hypothyroidism based on only the (autoimmune) hypothyroidism reports and separately list thyroiditis as adverse event (9, 18, 19). Besides the influence of the identification and subclassification, also other factors could be of influence on the reported incidence of thyroid disorders that were not assessed within this study. For example, the setting in which the study is conducted plays an important role. Mannesse et al. showed that the prevalence of hyponatremia was higher in patients admitted to geriatric wards as compared to other settings, e.g., nursing homes (20). Closely related to this, is the potential difference between the clinical trial setting and post-marketing setting. Recently, for the sodium-glucose cotransporter 2 inhibitor empagliflozin it was shown that approximately 55% of the clinical practice population would not have been eligible for inclusion in the clinical trials (21). Also, the follow-up time may play a role although for the thyroid disorders this influence is limited considering that thyroid disorders occur shortly after initiation, mostly within the first months of treatment (22, 23).

Clinical trials and observational studies serve different purposes and have a different approach, the identification and classification of thyroid disorders are, however, expected to be in line. In the light of the recent interests of implementing real-world evidence to support regulatory decision making and thereby (partly) replacing randomized controlled trials, standardization of information is vital. Initiatives were employed to study the practicalities hereof with a focus on replicating the effectiveness outcomes of randomized controlled trials in clinical practice (24, 25). However, as shown in our study also the safety perspective should be considered, given that challenges are faced including the standardizing of the identification and classification of the adverse events.

Although we only illustrated the differences in identification and classification of adverse events by using a case study of thyroid disorders associated with PD-1/PD-L1 inhibitors, it is expected that the results could also be applicable to other types of adverse events and drugs. Thyroid disorders can objectively be identified according to laboratory values, although minor differences may be expected introduced by the measurement method, these are expected to be uniformly identified. Since this study illustrated that for these adverse events differences in identification

and classification exist, it is expected that this is also the case for other adverse events for which the identification and classification can be further complicated as subjective measures are used for this. It is therefore considered that differences in the identification and classification of adverse events impact the regulatory and clinical interpretation of the safety information as, for example, the safety information described in the product information is based on data from the clinical trials.

In conclusion, the identification and classification of thyroid disorders differed between and within clinical trials and observational studies. Efforts to improve the comparability of the identification and classification of adverse events should be taken to facilitate regulatory and clinical interpretation of the safety profile.

REFERENCES

1. Bonhoeffer J, Kohl K, Chen R, Duclos P, Heijbel H, Heininger U, et al. The Brighton Collaboration: addressing the need for standardized case definitions of adverse events following immunization (AEFI). Vaccine. 2002;21(3-4):298-302.
2. Woodworth T, Furst DE, Alten R, Bingham C, Yocum D, Sloan V, et al. Standardizing assessment and reporting of adverse effects in rheumatology clinical trials II: the Rheumatology Common Toxicity Criteria v.2.0. The Journal of Rheumatology. 2007;34(6):1401-14.
3. Trotti A, Colevas AD, Setser A, Rusch V, Jaques D, Budach V, et al. CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. Semin Radiat Oncol. 2003;13(3):176-81.
4. Danan G, Benichou C. Causality assessment of adverse reactions to drugs—I. A novel method based on the conclusions of international consensus meetings: Application to drug-induced liver injuries. Journal of Clinical Epidemiology. 1993;46(11):1323-30.
5. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. N Engl J Med. 2020.
6. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. N Engl J Med. 2020;383(27):2603-15.
7. Voysey M, Clemens SAC, Madhi SA, Weckx LY, Folegatti PM, Aley PK, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. Lancet. 2021;397(10269):99-111.
8. Schuerch M, Gasse C, Robinson NJ, Alvarez Y, Walls R, Mors O, et al. Impact of varying outcomes and definitions of suicidality on the associations of antiepileptic drugs and suicidality: comparisons from UK Clinical Practice Research Datalink (CPRD) and Danish national registries (DNR). Pharmacoepidemiol Drug Saf. 2016;25 Suppl 1:142-55.
9. European Medicines Agency. Opdivo: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/opdivo-epar-product-information_en.pdf. Accessed 1/12/2020.
10. European Medicines Agency. Keytruda: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/keytruda-epar-product-information_en.pdf. Accessed 1/12/2020.
11. Onoi K, Chihara Y, Uchino J, Shimamoto T, Morimoto Y, Iwasaku M, et al. Immune Checkpoint Inhibitors for Lung Cancer Treatment: A Review. J Clin Med. 2020;9(5).
12. Sanlorenzo M, Vujic I, Posch C, Dajee A, Yen A, Kim S, et al. Melanoma immunotherapy. Cancer Biol Ther. 2014;15(6):665-74.
13. Postow MA, Sidlow R, Hellmann MD. Immune-Related Adverse Events Associated with Immune Checkpoint Blockade. N Engl J Med. 2018;378(2):158-68.
14. National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE) v5.0. https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_5x7.pdf. Accessed 8/1/2021.
15. Haanen J, Carbonnel F, Robert C, Kerr KM, Peters S, Larkin J, et al. Management of toxicities from immunotherapy: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2018;29(Suppl 4):iv264-iv6.
16. Brahmer JR, Lacchetti C, Schneider BJ, Atkins MB, Brassil KJ, Caterino JM, et al. Management of Immune-Related Adverse Events in Patients Treated With Immune Checkpoint Inhibitor Therapy: American Society of Clinical Oncology Clinical Practice Guideline. J Clin Oncol. 2018;36(17):1714-68.
17. European Medicines Agency. Tecentriq: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/tecentriq-epar-product-information_en.pdf. Accessed 1/12/2020.

18. European Medicines Agency. Imfinzi: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/imfinzi-epar-product-information_en.pdf. Accessed 2/12/2020.

19. European Medicines Agency. Bavencio: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/bavencio-epar-product-information_en.pdf. Accessed 5/2/2021.

20. Mannesse CK, Vondeling AM, van Marum RJ, van Solinge WW, Egberts TCG, Jansen PAF. Prevalence of hyponatremia on geriatric wards compared to other settings over four decades: A systematic review. Ageing Research Reviews. 2013;12(1):165-73.

21. Munk NE, Knudsen JS, Pottegård A, Witte DR, Thomsen RW. Differences Between Randomized Clinical Trial Participants and Real-World Empagliflozin Users and the Changes in Their Glycated Hemoglobin Levels. JAMA Netw Open. 2020;3(2):e1920949.

22. Olsson-Brown A, Lord R, Sacco J, Wagg J, Coles M, Pirmohamed M. Two distinct clinical patterns of checkpoint inhibitor-induced thyroid dysfunction. Endocrine connections. 2020;9(4):318-25.

23. Sakakida T, Ishikawa T, Uchino J, Chihara Y, Komori S, Asai J, et al. Clinical features of immune-related thyroid dysfunction and its association with outcomes in patients with advanced malignancies treated by PD-1 blockade. Oncology letters. 2019;18(2):2140-7.

24. Franklin JM, Patorno E, Desai RJ, Glynn RJ, Martin D, Quinto K, et al. Emulating Randomized Clinical Trials With Nonrandomized Real-World Evidence Studies: First Results From the RCT DUPLICATE Initiative. Circulation. 2021;143(10):1002-13.

25. Franklin JM, Pawar A, Martin D, Glynn RJ, Levenson M, Temple R, et al. Nonrandomized Real-World Evidence to Support Regulatory Decision Making: Process for a Randomized Trial Replication Project. Clin Pharmacol Ther. 2020;107(4):817-26.

SUPPLEMENTARY INFORMATION 1. PUBMED SEARCH STRATEGY

PD-1 inhibitor* [TIAB] OR PD1 inhibitor*[TIAB] OR PDL1 inhibitor*[TIAB] OR PD-L1 inhibitor*[TIAB] OR PD-1/PD-L1 inhibitor*[TIAB] OR PD-1/L1 inhibitor[TIAB] OR Checkpoint inhibitor*[TIAB] OR Checkpointinhibitor*[TIAB] OR anti-PD-1/PD-L1* OR Programmed Cell Death 1 [TIAB] * OR Nivolumab[TIAB] OR Pembrolizumab[TIAB] OR Atezolizumab[TIAB] OR Bavencio[TIAB] OR Avelumab[TIAB] OR Imfinzi [TIAB] OR Durvalumab[TIAB] OR Keytruda[TIAB] OR Pembrolizumab[TIAB] OR Libtayo[TIAB] OR Cemiplimab[TIAB] OR Opdivo[TIAB] OR Tecentriq [TIAB] OR Atezolizumab[TIAB] OR Immunotherapy [TIAB] OR Programmed Cell Death 1 Receptor [MeSH] OR pembrolizumab [Supplementary Concept] OR atezolizumab [Supplementary Concept] OR durvalumab [Supplementary Concept] OR nivolumab [MeSH] OR avelumab [Supplementary Concept] OR cemiplimab [Supplementary Concept]

4.1

AND

Thyroid Diseases/chemically induced[MeSH] OR Thyroid Diseases/physiopathology [MeSH] OR Thyroid Function Tests [MeSH] OR thyroid function test* [TIAB] OR thyroid disease* [TIAB] OR Thyroid abnormalit*[TIAB] OR Thyroid dysfunction*[TIAB] OR Hypothyroid* [TIAB] OR Hyperthyroid*[TIAB] OR Thyroiditis[TIAB] OR Abnormal thyroid function[TIAB] OR Thyroid* [TIAB] OR TSH [TIAB] OR FT4 [TIAB] OR thyroid gland [TIAB] OR thyroxine OR Thyrotoxicosis [TIAB] OR Thyroid stimulating hormone [TIAB] OR TSH [TIAB] OR Hyperthyroidism [MeSH] OR Thyroid Crisis [MeSH] OR Thyrotoxicosis [MeSH] OR Hypothyroidism [MeSH] OR Thyroid Stimulating Hormone Deficiency [MeSH] OR Thyroid-Stimulating Hormone Deficiencies [MeSH] OR Thyroid-Stimulating Hormone Deficiencies [TIAB] OR Thyroid-Stimulating Hormone Deficiency OR Thyroiditis, Autoimmune"[MeSH] OR thyroid diseases

SUPPLEMENTARY INFORMATION 2. CLINICAL STUDIES

Table S1. Characteristics of the main clinical studies supporting initial approval and extensions of indication for atezolizumab.

Study (NCT number)	Study design	Treatment schedule	Types of malignancy	# patients included	Follow-up time	Time frame
OAK (NCT02008227)	Randomized, open-label, phase 3 study	Atezolizumab 1200 mg, Q3W	Non–small cell lung cancer (NSCLC) after failure with platinum-containing chemotherapy	Randomized: 1225 Safety population atezolizumab: 609	Median follow-up: 32 months Median treatment duration: 3.4 months (range: 0 – 26)	Enrollment period: 11 March 2014 – 28 November 2014 (primary analysis) additional patients were enrolled until 29 April 2015
POPLAR (NCT01903993)	Randomized, open-label, phase 2 study	Atezolizumab 1200 mg, Q3W	NSCLC after failure with platinum-containing chemotherapy	Randomized: 287 Safety population atezolizumab: 142	Median follow-up: 14.8 months (range: 0.2+ – 19.6) Median treatment duration: 3.7 months (range: 0 – 19)	Enrollment period: 5 August 2013 – 31 March 2014
IMvigor 210 cohort 1 (NCT02108652)	Single-arm, open-label phase 2 study	Atezolizumab 1200 mg, Q3W	Cisplatin-ineligible locally-advanced or metastatic urothelial carcinoma (mUC)	Safety population atezolizumab: 123	Median follow-up: 17.2 months (range: 0.2 – 23.5) Median treatment duration: 15 weeks (range: 0 – 102)	Enrollment period: 9 June 2014 – 30 March 2015
IMvigor 210 cohort 2 (NCT02108652)	Single-arm, open-label, phase 2 study	Atezolizumab 1200 mg, Q3W	Inoperable locally advanced or mUC whose disease had progressed after prior platinum-based chemotherapy	Safety population atezolizumab: 315	Median follow-up: 11.7 months Median treatment duration: 12 weeks (range: 0 – 66)	Enrollment period: May 2014 – November 2014

Table S1. (continued)

Study (NCT number)	Study design	Treatment schedule	Types of malignancy	# patients included	Follow-up time	Time frame
IMvigor 211 (NCT02302807)	Randomized, open-label, phase 3 study	Atezolizumab 1200 mg, Q3W	Locally advanced or mUC who had progressed after platinum-based chemotherapy	Randomized: 931 Safety population atezolizumab: 459	Median follow-up: 17.3 months (range: 0 – 24.5) Median treatment duration: 2.8 months (range: 0 – 24)	Enrollment period: 13 January 2015 – 15 February 2016
IMpower150 (NCT02366143)	Randomized, open-label, phase 3 study	Atezolizumab 1200 mg Q3W + carboplatin + paclitaxel (ACP) Atezolizumab 1200 mg Q3W + bevacizumab + carboplatin + paclitaxel (ABCP)	Metastatic NSCLC who had not previously received chemotherapy	Randomized: 1202 Safety population atezolizumab: 393	Median follow-up: 15.4 months Median treatment duration: 8.2 months (range: 0 – 26) (ABCP group)	Enrollment period: March 2015 – December 2016
IMpower133 (NCT02763579)	Randomized, double blind, placebo controlled, phase 3 study	Atezolizumab 1200 mg Q3W + carboplatin + etoposide	Extensive-stage small-cell lung cancer who had not previously received treatment	Randomized: 403 Safety population atezolizumab: 198	Median follow-up: 13.9 months Median treatment duration: 4.7 months (range: 0 – 21)	Enrollment period: 6 June 2016 – 31 May 2017
IMpower 130 (NCT02367781)	Randomized, open-label, phase 3 study	Atezolizumab 1200 mg Q3W + carboplatin + nab-paclitaxel	Non-squamous non-small-cell lung cancer who received no previous chemotherapy for stage IV disease	Randomized: 724 Safety population atezolizumab: 473	Median follow-up: 18.5 months (IQR: 15.2 –23.6) Median treatment duration: 8.9 months (SD: 7.2)	Enrollment period: 16 April 2015 – 13 February 2017

Table S2. Description of the applied definitions to identify thyroid-related (immune-related) adverse events, frequency of screening of the thyroid levels and incidence of thyroid disorders in the clinical studies supporting initial approval and extensions of indication for atezolizumab.

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders	Definition used to identify immune-related adverse events
OAK (NCT02008227)		The incidence, nature, and severity of adverse events and laboratory abnormalities were assessed by the National Cancer Institute Common Terminology Criteria for Adverse Events version 4.0 (1)	Pooled safety population NSCLC + UC (2): Hyperthyroidism 1.7% Hypothyroidism 4.7% Pooled safety population NSCLC (3): Hypothyroidism/increased TSH: 42 (4.1%)	
POPLAR (NCT01903993)	Thyroid function testing including TSH (obtain free T3 and free T4 if abnormal result) at baseline and end-of study (3)	We graded adverse events with the National Cancer Institute Common Terminology Criteria for Adverse Events version 4.0 (4)	Hypothyroidism: 6% (3)	Immune-mediated adverse events: Patients who required the use of systemic corticosteroids within 30 days after the AE onset date (based on the concomitant medication CRF) with no clear alternate etiology. Systemic corticosteroids specifically excluded steroids administered via the following routes: inhaled, intranasal, intravitreal, ophthalmic, otic, per vagina, and topica (3)
IMvigor 210 cohort 1 (NCT02108652)	Thyroid function testing including TSH (obtain free T3 and free T4 if abnormal result) at baseline and end-of study (5)	Safety was assessed per National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (6)	Treatment-related adverse events: Hypothyroidism: 8 (7%) Adverse events regardless of attribution: Hypothyroidism: 9 (8%) Immune-mediated adverse events: Hypothyroidism: 1 (1%) (6)	Immune-mediated adverse events: The date of systemic corticosteroid initiation was on or up to 30 days after the AE onset date, the date of corticosteroid initiation was prior to the AE resolution date, and no clear alternate etiology could be identified (5)

Table S2. (continued)

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders	Definition used to identify immune-related adverse events
IMvigor 210 cohort 2 (NCT02108652)	Thyroid function testing including TSH (obtain free T3 and free T4 if abnormal result) at baseline and end-of study (5)	Safety assessments were performed according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE), Version 4.0 (7)	Hypothyroidism/increased TSH: 7 (2.3%) (5)	In line with IMvigor 210 cohort 1 (5)
IMvigor 211 (NCT02302807)	TSH, free T3, free T4: At baseline and on day 1 of cycles 1-5 and every four cycles thereafter and at treatment discontinuation (8)	National Cancer Institute Common Terminology Criteria for Adverse Events version 4.0 was used to assess adverse event frequency and severity (8)	Serious adverse events: Hyperthyroidism: 1 (0.22%) Hypothyroidism (serious adverse event): 2 (0.44%) (9) Pooled UC population: Hypothyroidism: 13 (2.5%) (5)	
IMpower150 (NCT02366143)	Thyroid function testing including thyroid-stimulating hormone, free T3, free T4 at baseline and every fourth cycle thereafter and at the treatment discontinuation visit (10)	Adverse events were assessed according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (10)	Immune-related adverse events: Hyperthyroidism: 16 (4.1%) Hypothyroidism: 50 (12.7%) (10)	Immune-related AEs were defined using MedDRA Preferred Terms that included both diagnosed immune conditions and signs and symptoms potentially representative of immune-related events, regardless of investigator-assessed causality (10)

Table S2. (continued)

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders	Definition used to identify immune-related adverse events
IMpower133 (NCT02763579)	Thyroid function testing (TSH, free T3, free T4) collected on Day 1 of Cycles 1, 4, 8, and 12, and every fourth cycle thereafter and at the treatment discontinuation visit (11)	Adverse events were assessed according to National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0. The investigators determined whether adverse events were related to the trial regimen (11)	Immune-related adverse events: Hyperthyroidism: 11 (5.6%) Hypothyroidism: 25 (12.6%) Serious treatment-related adverse events: Autoimmune thyroiditis: 2 (1%) (11)	In line with Impower150 (11)
IMpower 130 (NCT02367781)	Thyroid-function testing (thyroid-stimulating hormone [TSH], free T3, free T4) collected at Cycle 1, Day 1, and every fourth cycle thereafter and at the treatment discontinuation visit (12)	Incidence, nature, and severity of adverse events were graded in accordance with National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (12)	Adverse events of special interest: Hypothyroidism: 53 (11.2%) All-cause adverse events: Hyperthyroidism: 23 (4.9%) Hypothyroidism: 70 (14.8%) Blood thyroid stimulating hormone increased: 7 (1.5%) (12)	

Table S3. Characteristics of the main clinical studies supporting initial approval and extensions of indication for avelumab.

Study (NCT number)	Study design	Treatment schedule	Types of malignancy	# patients included	Follow-up time	Time frame
JAVELIN Merkel 200 (NCT02155647)	Single arm, open-label, phase 2 study	Avelumab, 10 mg/kg Q2W	Merkel cell carcinoma	Safety population avelumab: 88	Median follow-up: 10.4 months (IQR: 8.6 – 13.1) Median treatment duration: 17 weeks (IQR: 7 – 37)	Enrollment period: 25 July 2014 – 3 September 2015
JAVELIN Renal 101 (NCT02684006)	Randomized, open-label, phase 3 study	Avelumab, 10 mg/kg Q2W + axitinib	Advanced renal cell carcinoma	Randomized: 886 Safety population avelumab: 434	Median follow-up: 9.9 months Median treatment duration: 8.6 months (range: 0.5 – 25.3)	Enrollment period: 29 March 2016 – 19 December 2017

Table S4. Description of the applied definitions to identify thyroid-related (immune-related) adverse events, frequency of screening of the thyroid levels and incidence of thyroid disorders in the clinical studies supporting initial approval and extension of indication for avelumab.

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders	Definition used to identify immune-related adverse events
JAVELIN Merkel 200 (NCT02155647)	Free thyroxine (free T4), and thyroid stimulating hormone (TSH): at screening, every 3 months on treatment, and as medically indicated (13)	Safety was assessed according to the National Cancer Institute Common Terminology Criteria for Adverse Events version 4.0 (14)	Potential immune-mediated treatment-related adverse event: Hyperthyroidism: 2 (2.3%) Hypothyroidism: 3 (3.4%) (14)	Immune-mediated adverse events (imAEs): A two-level case definition was used: **Level 1:** A MedDRA PT query was established: *HLT thyroid hypofunction disorders:* autoimmune hypothyroidism, hypothyroidic goiter, hypothyroidism, myxoedema, primary hypothyroidism, secondary hypothyroidism, tertiary hypothyroidism, thyroid atrophy, transient hypothyroxinaemia or prematurity *HLT thyroid hyperfunction disorders:* Basedow's disease, hyperthyroidism, Marine Lenhart syndrome, primary hyperthyroidism, secondary hyperthyroidism, thyroid dermatopathy, thyrotoxic crisis, thyrotoxic periodic paralysis, toxic goitre, toxic nodular goiter *HLT acute and chronic thyroiditis:* autoimmune thyroiditis, thyroiditis, thyroiditis acute, thyroiditis chronic, thyroiditis fibrous chronic, thyroiditis subacute **Level 2:** AEs identified by the MedDRA PT queries were then medically reviewed using predefined case definitions for immune-related adverse reactions. All potential imARs were reviewed by 2 medically-qualified persons. If the 2 persons came to different assessments for a potential irAE, a third medically-qualified reviewer was asked to make the final assessment. The following criteria were used by the medical reviewers: *Onset:* AE onset after first avelumab administration until up to 90 days after last dose (end of AE collection period in the studies) *Duration:* AE does not spontaneously resolve (i.e., without corticosteroids / immunosuppressant treatment) within 7 days after onset *Immunosuppressive therapy*: AE treated with corticosteroid or other immunosuppressant therapy. For endocrinopathies only: AE required hormone replacement and / or (corticosteroid or other immunosuppressive therapy) *Etiology:* No other clear etiology or histopathology/biopsy consistent with immune-related event. (13)

Table S4. (continued)

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders	Definition used to identify immune-related adverse events
JAVELIN Renal 101 (NCT02684006)	Free T4 and TSH: at screening, lead-in Day 1 (only if not performed in prior 3 days), cycle 4 Day 1, cycle 7 Day 1, cycle 13 Day 1, Q12W thereafter, at end of treatment/ withdrawal, and 30-, 60-, and 90-days post-treatment. Additional tests should be performed when clinically indicated. (15)	Adverse events were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.03 (15)	Adverse events of any grade that occurred during treatment: Hypothyroidism: 108 (24.9%) Treatment-related adverse Events: Hypothyroidism: 105 (24.2) Immune-related thyroid disorders: 107 (24.7%) (15)	Immune-related adverse events: Event selected based on a list of prespecified MedDRA PTs within clusters (see JAVELIN Merkel 200). AE onset during 1st study drug administration or anytime thereafter through 90 days after last dose of study treatment. AE treated with corticosteroids or other immunosuppressant therapy. For endocrinopathies only: AE required hormone replacement. No clear etiology (other than immune mediated etiology) Histopathology/biopsy consistent with immune-mediated event. The dataset associated with irAEs may be refined based on medical review. (15)

Table S5. Characteristics of the main clinical study supporting initial approval for cemiplimab.

Study (NCT number)	Study design	Treatment schedule	Types of malignancy	# patients included	Follow-up time	Time frame
R2810-ONC-1540 (NCT02760498)	Single arm, phase 2 study	Cemiplimab, 2 mg/kg Q2W Cemiplimab, 350 mg Q3W	Locally advanced or metastatic cutaneous squamous-cell carcinoma (CSCC)	Safety population cemiplimab: Group 1 (weight-based dosing, mCSCC): 59 Group 2 (weight-based dosing, loCSCC): 78 Group 3 (fixed dose, mCSCC): 56	Group 1: Median follow-up time: 7.9 months Group 2: Median follow-up time: 9.3 months (IQR: 5.1 – 15.7) Group 3: Median follow-up time: 8.1 months (range: 0.6 – 14.1)	Enrollment period: Group 1: May 2016 – April 2017 Group 2: June 2016 – April 2018 Group 3: July 2017 – March 2018

Table S6. Description of the applied definitions to identify thyroid-related (immune-related) adverse events, frequency of screening of the thyroid levels and incidence of thyroid disorders in the clinical study supporting initial approval for cemiplimab.

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders	Definition used to identify immune-related adverse events
R2810-ONC-1540 (NCT02760498)	TSH: at treatment initiation, every cycle and at the end of study (16)	The severity of treatment-emergent adverse events was graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (version 4.03) (16, 17)	**Group 1:** Adverse events regardless of attribution: Hypothyroidism: 5 (8.5%) Hyperthyroidism: 1 (1.7%) (16) Investigator-assessed treatment-related adverse Events: Hypothyroidism: 4 (6.8%) Hyperthyroidism: 1 (1.7%) **Group 2:** Treatment-emergent adverse events regardless of attribution: Hypothyroidism: 10 (8%) (18) **Group 3:** Treatment-related adverse events: Hypothyroidism: 6 (10.7%) Immune-related adverse events: Hypothyroidism: 6 (10.7%) (17)	All AEs of unknown etiology associated with drug exposure should be evaluated to determine possible immune etiology. If an irAE is suspected, efforts should be made to rule out neoplastic, infectious, metabolic, toxin or other etiologic causes prior to labeling an AE as an irAE. (16)

Table S7. Characteristics of the main clinical studies supporting initial approval and extension of indication for durvalumab.

Study (NCT number)	Study design	Treatment schedule	Types of malignancy	# patients included	Follow-up time	Time frame
PACIFIC (NCT02125461)	Randomized, double-blind, placebo-controlled, phase 3 study	Durvalumab, 10 mg/kg Q2W	Locally advanced, unresectable, non–small-cell lung cancer	Randomized: 713 Safety population durvalumab: 475	Median follow-up: 14.5 months (range: 0.2 – 29.9) Median number of infusions: 20 (range: 1 – 27)	Enrollment period: May 2014 - April 2016
CASPIAN (NCT03043872)	Open-label, randomized, phase 3 study	Durvalumab, 1500 mg Q3W + platinum + etoposide for up to four cycles followed by durvalumab 1500 mg Q4W Durvalumab, 1500 mg Q3W + tremelimumab + platinum + etoposide for up to four cycles followed by durvalumab 1500 mg Q4W	Extensive-stage small cell lung cancer	Randomized: 805 Safety population durvalumab: 265	Median follow-up: 14.2 months (IQR: 11.7 –17.0) Median number of doses: 7 (IQR: 6 – 11)	Enrollment period: 27 March 2017 – 29 May 2018

Table S8. Description of the applied definitions to identify thyroid-related (immune-related) adverse events, frequency of screening of the thyroid levels and incidence of thyroid disorders in the main clinical studies supporting initial approval and extension of indication for durvalumab.

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders	Definition used to identify immune-related adverse events
PACIFIC (NCT02125461)	Thyroid function tests (TSH and T3 and T4): at screening, randomization, every 2 weeks on treatment and 30 days after last dose (19)	All toxicities will be graded according to CTCAE Version 4.03 (19)	Adverse events of any cause: Hypothyroidism: 55/475 (11.6%) Treatment-related adverse events: Hypothyroidism: 50 (10.5%) Hyperthyroidism: 30 (6.3%) Immune-related adverse events: Hypothyroidism: 44 (9.3%) Hyperthyroidism: 13 (2.7%) (19)	An adverse event of special interest requiring the use of systemic steroids or other immunosuppressants, and/or, for specific endocrine events, endocrine therapy, consistent with an immune-mediated mechanism of action, and where there is no clear alternate etiology. (19)

Table S8. (continued)

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders	Definition used to identify immune-related adverse events
CASPIAN (NCT03043872)	TSH: at screening, every cycle, 28 days, 2 and 3 months following the last dose (20)	Adverse events were graded according to National Cancer Institute common terminology criteria for adverse events, version 4.03 (21)	**Durvalumab + tremelimumab + platinum + etoposide:** Adverse events of any cause: Hyperthyroidism: 28 (11%) Treatment-related adverse events as assessed per the investigator: Hyperthyroidism: 24 (9%) Hypothyroidism: 22 (8%) Immune-mediated adverse events: Hypothyroid events: 24 (9%) Hyperthyroid events: 21 (8%) **Durvalumab + platinum + etoposide:** Adverse events of any cause: Hyperthyroidism: 26 (10%) Treatment-related adverse events as assessed per the investigator: Hyperthyroidism: 22 (8%) Hypothyroidism: 23 (9%) Immune-mediated adverse events: Hypothyroid events: 24 (9%) Hyperthyroid events: 14 (5%) (21)	An immune-mediated adverse event is defined as an event that is associated with drug exposure and consistent with an immune-mediated mechanism of action, where there is no clear alternate aetiology and the event required treatment with systemic corticosteroids or other immunosuppressants and/or, for specific endocrine events, endocrine therapy (21)

Table S9. Characteristics of the main clinical studies supporting initial approval and extensions of indication for nivolumab.

Study (NCT number)	Study design	Treatment schedule	Types of malignancy
CheckMate-066 (NCT01721772)	Randomized, double-blind, phase 3 study	Nivolumab, 3 mg/kg Q2W	Unresectable or metastatic melanoma
CheckMate-037 (NCT01721746)	Randomized, open-label, phase 3 study	Nivolumab, 3 mg/kg Q2W	Unresectable or metastatic melanoma
CheckMate-017 (NCT01642004)	Randomized, open-label, phase 3 study	Nivolumab, 3 mg/kg Q2W	Advanced squamous-cell non–small-cell lung cancer (NSCLC)
CheckMate-057 (NCT01673867)	Randomized, open-label, phase 3 study	Nivolumab, 3 mg/kg Q2W	Metastatic nonsquamous NSCLC
CheckMate-025 (NCT01668784)	Randomized, open-label, phase 3 study	Nivolumab, 3 mg/kg Q2W	Advanced or metastatic clear-cell renal cell carcinoma
CheckMate-067 (NCT01844505)	Randomized, double-blind, phase 3 study	Nivolumab, 3 mg/kg Q2W Nivolumab 1 mg/kg Q3W + ipilimumab 3 mg/kg Q3W for 4 doses, followed by 3 mg/kg nivolumab Q2W	Unresectable or metastatic melanoma
CheckMate-205 (NCT02181738)	Single-arm, multi-cohort, open-label, phase 2 study	Nivolumab, 3 mg/kg Q2W	Classical Hodgkin Lymphoma

# patients included	Follow-up time	Time frame
Randomized: 418 Safety population nivolumab: 206	Follow-up: up to 16.7 months (database lock: 5.2 months after the first visit of the last patient who had undergone randomization)	Enrollment period: January 2013 – February 2014
Randomized: 405 Safety population nivolumab: 268	Median follow-up: 8.4 months (IQR: 7.0 – 9.8) Median treatment duration: 5.3 months (95% CI: 3.3 – 6.5)	Enrollment period: 21 December 2012 – 10 January 2014
Randomized: 272 Safety population nivolumab: 131	Minimum follow-up: approximately 11 months Median number of doses: 8 (range: 1 – 48)	Enrollment period: October 2012 – December 2013
Randomized: 582 Safety population nivolumab: 287	Minimum follow-up: 13.2 months Median number of doses: 6 (range: 1 – 52)	Enrollment period: November 2012 – December 2013
Randomized: 821 Safety population nivolumab: 406	Minimum follow-up: 14 months Median treatment duration: 5.5 months (range: <0.1 – 29.6)	Enrollment period: October 2012 – March 2014
Randomized: Safety population nivolumab: 313 (nivolumab monotherapy), 313 (nivolumab + ipilimumab)	Median follow-up: Ranging from 12.2 to 12.5 months Median number of doses: 15 (range: 1 – 38) (nivolumab monotherapy) 4 (range: 1 – 39) (nivolumab + ipilimumab)	Enrollment period: July 2013 – March 2014
Enrolled: 63 (brentuximab vedotin (BV)-naïve) 80 (BV received after auto-HCT) 100 (BV received before and/or after auto-HCT)	Median follow-up: 18 months	Enrollment period: August 2014 – August 2015

Table S9. (continued)

Study (NCT number)	Study design	Treatment schedule	Types of malignancy
CheckMate-141 (NCT02105636)	Open-label, randomized, phase 3 study	Nivolumab, 3 mg/kg Q2W	Metastatic Platinum-Refractory Squamous Cell Carcinoma of the Head and Neck
CheckMate-275 (NCT02387996)	Single-arm, phase 2 study	Nivolumab, 3 mg/kg Q2W	Metastatic or surgically unresectable locally advanced urothelial carcinoma
CheckMate-238 (NCT02388906)	Randomized, double-blind, phase 3 study	Nivolumab, 3 mg/kg Q2W	Melanoma with involvement of lymph nodes or metastatic disease who have undergone complete resection
CheckMate-214 (NCT02231749)	Randomized, open-label, phase 3 study	Nivolumab 3 mg/kg + ipilimumab 1 mg/kg Q3W for 4 doses followed by nivolumab 3 mg/kg Q2W	Previously untreated advanced renal cell carcinoma
CheckMate-9LA (NCT03215706)	Randomized, open-label, phase 3 study	Nivolumab, 360 mg Q3W + ipilimumab, 1 mg/kg Q6W + chemotherapy (2 cycles)	Stage IV NSCLC

Table S10. Description of the applied definitions to identify thyroid-related (immune-related) adverse events, frequency of screening of the thyroid levels and incidence of thyroid disorders in the clinical studies supporting initial approval and extensions of indication for nivolumab.

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders
CheckMate-066 (NCT01721772)	TSH, free T4, free T3: at screening. TSH (with reflexive free T4/free T3): every 6 weeks, 30 days from last dose (repeat at 100-121 days from last dose if study drug related toxicity persists) (22)	The severity of adverse events was graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (22)

# patients included	Follow-up time	Time frame
Randomized: 240 Safety population nivolumab: 236	Median follow-up: 5.1 months (range: 0 – 16.8) Median treatment duration: 1.9 months	Enrollment period: June 2014 – August 2015
Safety population nivolumab: 270	Median follow-up: 7 months (IQR: 2.96 – 8.77)	Enrollment period: 9 March 2015 – 16 October 2015
Randomized: 906 Safety population nivolumab: 452	Median follow-up: 19.5 months Median number of doses: 24 (range: 1 –26)	Enrollment period: 30 March 2015 – 30 November 2015
Randomized: 1096 Safety population nivolumab + ipilimumab: 547	Median follow-up: 25.2 months Median treatment duration: 7.9 months (95% CI, 6.5 – 8.4)	Enrollment period: October 2014 – February 2016
Randomized: 719 Safety population nivolumab + ipilimumab: 358	Median follow-up: 10.4 months Median treatment duration: 6.05 months (95% CI: 4.93 – 7.06)	NA

Incidence thyroid disorders	Definition used to identify immune-related adverse events
Treatment-related adverse events: Hypothyroidism: 9 (4.4%) Hyperthyroidism: 7 (3.4%) Select adverse events related to study treatment (those with potential immunological etiology): Hypothyroidism: 9 (4.4%) Hyperthyroidism: 7 (3.4%) (22)	

Table S10. (continued)

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders
CheckMate-037 (NCT01721746)		The site investigator graded adverse events according to the National Cancer Institute Common Terminology Criteria for Adverse Events version 4.0 throughout the study until 100 days after discontinuation of study treatment (23)
CheckMate-017 (NCT01642004)	TSH, free T4, free T3: at screening. TSH (with reflexive free T4/free T3): every 6 weeks, 30 days from last dose (repeat at 100-121 days from last dose if study drug related toxicity persists) (25)	Safety was assessed by means of evaluations of the incidence of adverse events, which were graded with the use of the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (25)
CheckMate-057 (NCT01673867)	TSH (reflex to free T4/free T3): at screening, every 6 weeks, 30 days from last dose (repeat at 100-121 days from last dose if study drug related toxicity persists) (27)	Safety was assessed by an evaluation of the incidence of clinical adverse events and laboratory variables, which were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (27)

Incidence thyroid disorders	Definition used to identify immune-related adverse events
Treatment-related select AEs of potential immune-related etiology: Hypothyroidism: 15 (5.6%) Blood TSH increased: 3 (1.1%) Hyperthyroidism: 5 (1.9%) (23) Endocrine adverse events: Thyroiditis (composite term which includes autoimmune thyroiditis): 1 (0.4%) Blood TSH decreased: 1 (0.4%) Thyroxine Free increased: 1 (0.4%) Abnormal thyroid tests: TSH > ULN: 59 (25%) With at least one FT3/FT4 test value <LLN: 26 (11.0%) With all other FT3/FT4 test value >=LLN: 24 (10.2%) TSH < LLN: 31 (13.1%) With at least one FT3/FT4 test value >ULN: 10 (4.2%) With all other FT3/FT4 test value <=ULN: 16 (6.8%) (24) Treatment-related select AEs: Hypothyroidism: 5 (4%) (25) Pooled safety population squamous NSCLC: Hypothyroidism: common Thyroiditis: uncommon (26) All causality AEs: Hypothyroidism: 19 (7%) Treatment-related Adverse Events: Hypothyroidism: 19 (7%) Treatment-related select adverse Events: Blood thyroid stimulating hormone increased: 6 (2%) Hyperthyroidism: 4 (1%) Blood thyroid stimulating hormone decreased: 1 (<1%) Thyroiditis: 1 (<1%) (27)	There were no case definitions for the identification of immune-mediated AEs in the protocol (24)

Table S10. (continued)

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders
Checkmate-025 (NCT01668784)	TSH (reflex to free T4/free T3): at screening, every 8 weeks, 30 days from last dose (repeat at 100-121 days from last dose if study drug related toxicity persists) (28)	Adverse events were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (28)
Checkmate-067 (NCT01844505)	TSH (reflex to free T4/free T3): at screening, every 3 weeks for the first 12 weeks followed by every 4 weeks, 30 days from last dose (repeat at 100-121 days from last dose if study drug related toxicity persists) (30)	The severity of adverse events was graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (30)
CheckMate-205 (NCT02181738)	TSH (reflex to free T4/free T3): at screening, every 6 weeks (31)	Adverse events were assessed using NCI CTCAE v. 4.0 (31)
CheckMate-141 (NCT02105636)	TSH (reflex to free T4/free T3): at screening, every other week (33)	At each treatment visit and for 100 days after receipt of the last dose, acute toxic effects were evaluated according to the Common Terminology Criteria for Adverse Events, version 4.0 (33)

Incidence thyroid disorders	Definition used to identify immune-related adverse events
Drug related select AEs: Hypothyroidism: 24 (5.9%) (28) Abnormal thyroid tests: TSH > ULN: 148 (38.7%) With at least one FT3/FT4 test value <LLN: 51 (13.4%) With all other FT3/FT4 test value >=LLN: 38 (9.9%) TSH < LLN: 59 (15.4%) With at least one FT3/FT4 test value >ULN: 19 (5.0%) With all other FT3/FT4 test value <=ULN: 19 (5.0%) (29)	
Nivolumab Treatment-related select adverse events: Hypothyroidism: 27 (8.6%) Hyperthyroidism: 13 (4.2%)	
Nivolumab + ipilimumab Treatment-related select adverse events: Hypothyroidism: 47 (15.0%) Hyperthyroidism: 31 (9.9%) (30)	
All-cause immune-mediated AEs Hypothyroidism: 21 (9%) Thyroiditis: 2 (<1%) Hyperthyroidism: 6 2%) (32)	All-cause immune-mediated AEs: include events defined as AEs (regardless of causality) that required immune-modulating medication (with the exception of those of endocrine origin) and were reported up to 100 days after the last dose (32)
Select treatment-related adverse events: Hypothyroidism: 9 (3.8%) Blood TSH increased: 3 (1.3%) Hyperthyroidism: 2 (0.8%) Abnormal thyroid function test: 2 (0.8%) Thyroiditis: 2 (0.8%) (33)	

Table S10. (continued)

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders
CheckMate-275 (NCT02387996)	TSH, free T4, free T3: at screening. TSH (with reflexive free T4/free T3): every other week, 35 (+/- 7 days), 80 (+/- 7 days) after last dose (34)	Adverse events were graded according to National Cancer Institute Common Terminology Criteria for Adverse Events (v4.0) during treatment and up to 30 days after treatment discontinuation (34)
CheckMate-238 (NCT02388906)	TSH, free T4, free T3: at screening. TSH (with reflexive free T4/free T3): every four weeks, 30 (+/- 7 days), 84 (+/- 7 days) after last dose (36)	Data regarding adverse events were collected for each group according to the Common Terminology Criteria for Adverse Events, version 4.0 (36)
CheckMate-214 (NCT02231749)	TSH, free T4, free T3: at screening. TSH (with reflexive free T4/free T3): every three weeks for the first 12 weeks followed by every 4 weeks, 30 (+/- 7 days), 84 (+/- 7 days) after last dose	Adverse events were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (37)
CheckMate-9LA (NCT03215706)	TSH, free T4, free T3: at screening. TSH (with reflexive free T4/free T3): every 6 weeks, 35 (+/- 7 days) after last dose (repeat at 115 (+/- 7 days) if study drug related toxicity persists) (39)	Descriptive statistics of safety will be presented using National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.0 (39)

Incidence thyroid disorders	Definition used to identify immune-related adverse events
Treatment-related adverse events: Hypothyroidism: 21 (7.8%) (34) Drug-related selected adverse events: Hyperthyroidism: 11 (4.1%) Blood thyroid stimulating hormone increased: 10 (3.7%) Blood thyroid stimulating hormone decreased: 5 (1.9%) Thyroiditis: 2 (0.7%) Thyroxine increased: 2 (0.7%) Autoimmune thyroiditis: 1 (0.4%) Thyroxine decreased: 1 (0.4%) Thyroxine free increased: 1 (0.4%) (35)	
Treatment-related adverse events: Hypothyroidism: 49 (10.8%) Hyperthyroidism: 36 (8.0) (36)	Immune-related selected adverse events were determined on the basis of a prespecified list of terms from the Medical Dictionary for Regulatory Activities, which was updated according to each new version. (36)
Treatment-related adverse events: Hypothyroidism: 85 (15.5%) (37) Drug-related selected adverse events: Hyperthyroidism: 59 (10.8%) Blood thyroid stimulating hormone increased: 11 (2.0%) Blood thyroid stimulating hormone decreased: 5 (0.9%) Thyroiditis: 16 (2.9%) Basedow's disease: 2 (0.4%) Thyroxine free increased: 2 (0.4%) Autoimmune hypothyroidism: 1 (0.2%) Autoimmune thyroiditis: 1 (0.2%) Thyroid function test abnormal: 1 (0.2%) Thyroxine decreased: 1 (0.2%) (38)	
Adverse events: Hyperthyroidism: 29 (8.1%) Hypothyroidism: 55 (15.4%) (39)	

Table S11. Characteristics of the main clinical studies supporting initial approval and extensions of indication for pembrolizumab.

Study (NCT number)	Study design	Treatment schedule	Types of malignancy
Keynote-002 (NCT01704287)	Randomized, open-label, phase 2 study	Pembrolizumab 2 mg/kg Q3W Pembrolizumab 10 mg/kg Q3W	Ipilimumab-refractory melanoma
Keynote-006 (NCT01866319)	Randomized, open-label, phase 3 study	Pembrolizumab 10 mg/kg Q2W Pembrolizumab 10 mg/kg Q3W	Advanced melanoma
Keynote-010 (NCT01905657)	Randomized, open-label, phase 2/3 study	Pembrolizumab 2 mg/kg Q3W Pembrolizumab 10 mg/kg Q3W	Advanced non-small-cell lung cancer
Keynote-024 (NCT02142738)	Randomized, open-label, phase 3 study	Pembrolizumab, 200 mg Q3W	Metastatic non-small cell lung carcinoma
Keynote-087 (NCT02453594)	Single arm, phase 2 study	Pembrolizumab, 200 mg Q3W	Relapsed or refractory classical Hodgkin lymphoma
Keynote-045 (NCT02256436)	Open-label, randomized, phase 3 study	Pembrolizumab, 200 mg Q3W	Locally advanced or metastatic urothelial carcinoma
Keynote-052 (NCT02335424)	Single arm, phase 2 study	Pembrolizumab, 200 mg Q3W	Locally advanced or metastatic urothelial carcinoma

# patients included	Follow-up time	Time frame
Randomized: 540 Safety population pembrolizumab: 180 (2 mg/kg group), 181 (10 mg/kg group)	Median follow-up: 10 months (IQR: 8 – 12) Median treatment duration: 113 days (range 1 – 499) (2 mg/kg group), 145 days (range: 1 – 505) (10 mg/kg group)	Enrollment period: 30 November 2012 – 13 November 2013
Randomized: 834 Safety population pembrolizumab: 278 (Q2W group), 277 (Q3W group)	Median follow-up: 7.9 months (range: 6.1 – 11.5) Median treatment duration: 164 days (Q2W group), 151 (Q3W group)	Enrollment period: 18 September 2013 – 3 March 2014
Randomized: 1034 Safety population pembrolizumab: 339 (2 mg/kg group), 343 (10 mg/kg group)	Median follow-up: 13.1 months (IQR: 8.6 –17.7) Median treatment duration: 3.5 months (IQR: 1.4 – 7.2) (2 mg/kg group), 3.5 months (1.4 – 7.0) (10 mg/kg group)	Enrollment period: 28 August 2013 – 27 February 2015
Randomized: 305 Safety population pembrolizumab: 154	Median follow-up: 11.2 months (range: 6.3 – 19.7) Median treatment duration: 7.0 months (range: 1 day – 18.7 months)	Enrollment period: 19 September 2014 – 29 October 2015
Safety population pembrolizumab: 210	Median follow-up: 10.1 months (range: 1.0 – 15.0) Median treatment duration: 8.3 months (range: 0.03 – 14.99)	Enrollment period: 26 June 2015 - 21 March 2016
Randomized: 542 Safety population pembrolizumab: 266	Median follow-up: 14.1 months (range: 9.9 – 22.1) Median treatment duration: 3.5 months (range: <0.2 – 20.0)	Enrollment period: 5 November 2014 – 13 November 2015
Safety population pembrolizumab: 370	Median follow-up: 5 months (IQR: 3.0 – 8.6) Median treatment duration: 3 months (range: 0.03 – 16.0)	Enrollment period: 24 February 2015 – 8 August 2016

Table S11. (continued)

Study (NCT number)	Study design	Treatment schedule	Types of malignancy
Keynote-040 (NCT02252042)	Randomized, open-label, phase 3 study	Pembrolizumab, 200 mg Q3W	Recurrent or metastatic head-and-neck squamous cell carcinoma
Keynote-189 (NCT02578680)	Randomized, double-blind, phase 3 study	Pembrolizumab, 200 mg Q3W + pemetrexed and a platinum-based drug	Metastatic nonsquamous NSCLC
Keynote-054 (NCT02362594)	Randomized, double-blind, placebo-controlled, phase 3 study	Pembrolizumab, 200 mg Q3W	Resected, high-risk stage III melanoma
Keynote-407 (NCT02775435)	Randomized, double-blind, placebo-controlled, phase 3 study	Pembrolizumab 200 mg Q3W + 4 cycles of carboplatin + paclitaxel/nab-paclitaxel	Metastatic, squamous NSCLC
Keytruda-426 (NCT02853331)	Randomized, open-label, phase 3 study	Pembrolizumab, 200 mg Q3W + axitinib	Locally advanced or metastatic renal cell carcinoma
Keynote-048 (NCT02358031)	Randomized, open-label, phase 3 study	Pembrolizumab, 200 mg Q3W Pembrolizumab, 200 mg Q3W + platinum + 5-fluorouracil	Recurrent or metastatic squamous cell carcinoma of the head and neck

# patients included	Follow-up time	Time frame
Randomized: 495 Safety population pembrolizumab: 246	Median follow-up: 7.5 months (IQR: 3.4 – 13.3) Median treatment duration: 2.8 months (IQR: 1.2 – 6.8)	Enrollment period: 24 December 2014 – 13 May 2016
Randomized: 616 Safety population pembrolizumab: 405	Median follow-up: 10.5 months (range: 0.2 – 20.4) Mean treatment duration: 7.4 months (SD: 4.7)	Enrollment period: 26 February 2016 – 6 March 2017
Randomized: 1019 Safety population pembrolizumab: 509	Median follow-up: 14.7 months Median number of doses: 18 (IQR: 8 – 18)	Enrollment period: August 2015 – November 2016
Randomized: 559 Safety population pembrolizumab: 278	Median follow-up: 7.8 months (range: 0.1 – 19.1) Mean treatment duration: 6.3 months (SD: 4.1)	Enrollment period: 19 August 2016 – 28 December 28 2017
Randomized: 861 Safety population pembrolizumab: 429	Median follow-up: 12.8 months (range: 0.1 – 22.0) Median treatment duration: 10.4 months (range: 0.03 – 21.2)	Enrollment period: 24 October 2016 – 24 January 2018
Randomized: 882 Safety population pembrolizumab: 300 (pembrolizumab alone) 276 (pembrolizumab with chemotherapy)	Median follow-up: 11.5 months (IQR: 5.1 – 20.8) (pembrolizumab alone group), 13.0 (IQR: 6.4 – 21.5) (pembrolizumab with chemotherapy group) Median treatment duration: 3.5 months (IQR: 1.4 – 7. 6) (pembrolizumab alone group), 5.8 months (IQR: 2.8 – 9.7) (pembrolizumab with chemotherapy group)	Enrollment period: 20 April 2015 – 17 January 2017

Table S12. Description of the applied definitions to identify thyroid-related (immune-related) adverse events, frequency of screening of the thyroid levels and incidence of thyroid disorders in the clinical studies supporting initial approval and extensions of indication for pembrolizumab.

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders
Keynote-002 (NCT01704287)	T3, FT4, TSH: at screening, from cycle 2 every 12 weeks, 30 days (+-3 days) following the last dose (40)	Adverse events, laboratory values, and vital signs were assessed regularly throughout the study and graded per the Common Terminology Criteria for Adverse Events, version 4.0 (40)
Keynote-006 (NCT01866319)	T3, FT4, TSH: Q2W group: week 0, from cycle 2 every other cycle, end of treatment visit (30 +-3 days following last dose) Q3W group: every cycle for the first four cycles followed by every other cycle, end of treatment visit (30 +-3 days following last dose) (41)	Adverse events, laboratory values, and vital signs were assessed regularly and graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (41)

Incidence thyroid disorders	Definition used to identify immune-related adverse events
2 mg/kg group: Treatment-related adverse events: Hypothyroidism: 9 (5%) Adverse events of a potentially immune-mediated nature, regardless of attribution: Hypothyroidism: 11 (6%) Hyperthyroidism: 7 (4%) **10 mg/kg group:** Treatment-related adverse events: Hypothyroidism: 13 (7%) Adverse events of a potentially immune-mediated nature, regardless of attribution: Hypothyroidism: 15 (8%) Hyperthyroidism: 2 (1%) (40)	An irAE may be defined as an adverse event of unknown etiology, associated with drug exposure and is consistent with an immune phenomenon. Efforts should be made to rule out neoplastic, infectious, metabolic, toxin or other etiologic causes prior to labeling an adverse event of clinical interest. Immunological, serological and histological (biopsy) data should be used to support the diagnosis of an immune-related toxicity. (40)
Q2W group: Adverse events of special interest on the basis of the likely autoimmune or immune-related mechanism: Hypothyroidism: 28 (10.1%) Hyperthyroidism: 18 (6.5%) Adverse Events Attributed to Study Treatment by the Investigator: Hypothyroidism: 25 (9.0%) Hyperthyroidism: 17 (6.1%) Blood thyroid stimulating hormone decreased: 3 (1.1%) Blood thyroid stimulating hormone increased: 2 (0.7%) **Q3W group:** Adverse events of special interest on the basis of the likely autoimmune or immune-related mechanism: Hypothyroidism: 24 (8.7%) Hyperthyroidism: 9 (3.2%) Adverse Events Attributed to Study Treatment by the Investigator: Hypothyroidism: 21 (7.6%) Hyperthyroidism: 7 (2.5%) Blood thyroid stimulating hormone decreased: 3 (1.1%) Blood thyroid stimulating hormone increased: 3 (1.1%) (41)	In line with Keynote-002 (41)

Table S12. (continued)

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders
Keynote-010 (NCT01905657)	T3/FT3, FT4 and TSH: at screening, every other cycle, and end of treatment visit (30 +-3 days following last dose) (42)	Adverse events were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (version 4.0) (42)
Keynote-024 (NCT02142738)	T3, free T4, and TSH: at screening, every other cycle, and end of treatment visit (30 +-3 days following last dose) (43)	All adverse events and abnormalities were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (43)
Keynote-087 (NCT02453594)	T3 (or FT3 per local standard), FT4 and TSH: at screening, every other cycle, and end of treatment visit (30 +-3 days following last dose) (44)	AEs were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (version 4.0) (44)

Incidence thyroid disorders	Definition used to identify immune-related adverse events
2 mg/kg group: Adverse events of special interest: Hypothyroidism: 28 (8%) Hyperthyroidism: 12 (4%) Thyroiditis: 2 (1%) Adverse Events Attributed to Study Treatment by the Investigator: Hypothyroidism: 25 (7.4%) Hyperthyroidism: 10 (2.9%) Blood thyroid stimulating hormone increased: 3 (0.9%) Thyroxine free increased: 1 (0.3%)	Subjects who develop an Event of clinical interest (including hypothyroidism, hyperthyroidism, thyroid disorder, thyroiditis) thought to be immune-related should have additional testing to rule out other etiologic causes. If lab results or symptoms indicated a possible immune-related ECI then additional testing should be performed to rule out other etiologic causes. If no other cause was found, then it is assumed to be immune-related. (42)
10 mg/kg group: Adverse events of special interest: Hypothyroidism: 28 (8%) Hyperthyroidism: 20 (6%) Thyroiditis: 0 (0%) Adverse Events Attributed to Study Treatment by the Investigator: Hypothyroidism: 23 (6.7%) Hyperthyroidism: 15 (4.4%) Blood thyroid stimulating hormone increased: 4 (1.2%) Thyroxine free increased: 4 (1.2%) (42)	
Immune-mediated adverse events (both that were and were not attributed to study treatment by the investigator): Hypothyroidism: 14 (9.1%) Hyperthyroidism: 12 (7.8%) Thyroiditis: 4 (2.6%) (43)	In line with Keynote-010 (43)
All-cause adverse events: Hypothyroidism: 29 (13.8%) Treatment-related adverse events: Hypothyroidism: 26 (12.4%) Immune-mediated adverse events (regardless of treatment attribution): Hypothyroidism: 29 (13.8%) Hyperthyroidism: 6 (2.9%) (44)	In line with Keynote-011 (44)

Table S12. (continued)

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders
Keynote-045 (NCT02256436)	T3, FT4 and TSH: at screening, every other cycle, and end of treatment visit (30 +-3 days following last dose) (45)	All the adverse events and abnormalities were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0. (45)
Keynote-052 (NCT02335424)	T3, FT4 and TSH: at screening, every other cycle, and end of treatment visit (30 +-3 days following last dose) (46)	Adverse events were monitored throughout the treatment period and for 30 days after treatment end (90 days for serious adverse events) and were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (version 4.0) (46)
Keynote-040 (NCT02252042)	T3, FT4 and TSH: at screening, every other cycle, and end of treatment visit (30 +-3 days following last dose) (48)	Adverse events and laboratory abnormalities were collected throughout treatment and for 30 days thereafter (90 days for serious adverse events and those of special interest to pembrolizumab treatment) and graded using the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (48)
Keynote-189 (NCT02578680)	T3 or FT3, FT4, and TSH: at screening, every other cycle, and end of treatment visit (30 +-3 days following last dose) (49)	Adverse events and laboratory abnormalities were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0. (49)

Incidence thyroid disorders	Definition used to identify immune-related adverse events
Adverse event of interest (with immune-related cause regardless of treatment attribution): Hypothyroidism: 17 (6.4%) Hyperthyroidism: 10 (3.8%) Thyroiditis: 2 (0.8%) (45)	In line with Keynote-011 (45)
Treatment-related adverse events: Thyroiditis: 2 (1%) Hypothyroidism: 42 (11.4%) (46)	In line with Keynote-011 (46)
Pooled Keynote-045/052: Drug-related adverse events Hyperthyroidism: 18 (2.8%) Hypothyroidism: 53 (8.3%) Blood thyroid stimulating hormone increased: 5 (0.8%) Thyroiditis: 3 (0.5%) (47)	
Treatment-related adverse events: Hypothyroidism: 33 (13.4%) Hyperthyroidism: 5 (2.0%) Blood thyroid stimulating hormone increased: 2 (0.8%) Thyroxine free increased: 1 (0.4%) Triiodothyronine decreased: 1 (0.4%) Adverse event of interest (with immune-related cause regardless of treatment attribution): Hypothyroidism: 37 (15%) Hyperthyroidism: 5 (2%) (48)	In line with Keynote-011 (48)
Pembrolizumab-pemetrexed-carboplatin group Adverse event of interest (with immune-related cause regardless of treatment attribution): Hypothyroidism: 20 (6.8%) Hyperthyroidism: 13 (4.4%)	In line with Keynote-011 (49)
Pembrolizumab-pemetrexed-cisplatin group: Adverse event of interest (with immune-related cause regardless of treatment attribution): Hypothyroidism: 7 (6.3%) Hyperthyroidism: 3 (2.7%) Thyroiditis: 1 (0.9%) (49)	

Table S12. (continued)

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders
Keynote-054 (NCT02362594)	TSH (in case of elevated TSH to add free T3 and T4): prior to randomization, every 6 weeks, 12 weeks (+- 2 weeks) after the last treatment administration. (50)	Data on adverse events were collected for each treatment course with the use of the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (50)
Keynote-407 (NCT02775435)	T3/FT3, FT4 and TSH: at screening, every other cycle, end of treatment visit (≤30 following last dose) (51)	Adverse events and abnormal laboratory findings were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.03 (51)
Keytruda-426 (NCT02853331)	T3/FT3, FT4, and TSH: at screening, every other cycle, end of treatment visit, and safety follow-up visit (30 +-3 days following last dose) (52)	Data on adverse events and laboratory abnormalities were collected regularly throughout the treatment period and for 30 days thereafter (data on serious adverse events and events of interest were collected for 90 days after the end of the treatment period) and were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0 (52)

Incidence thyroid disorders	Definition used to identify immune-related adverse events
Immune-related adverse events, regardless of investigator attribution: Hypothyroidism: 73 (14.3%) Hyperthyroidism: 52 (10.2%) Thyroiditis: 16 (3.1%) (50)	In line with Keynote-011 Immune-related adverse events were programmatically determined from a predefined list of *Medical Dictionary for Regulatory Activities* (MedDRA) terms, which was updated in accordance with each new version of MedDRA. (50)
Adverse event of interest (with immune-related cause regardless of treatment attribution): Hypothyroidism: 22 (7.9%) Hyperthyroidism: 20 (7.2%) Thyroiditis: 3 (1%) (51)	In line with Keynote-011 (51)
Adverse events of any cause: Hypothyroidism: 152 (35.4%) Hyperthyroidism: 55 (12.8%) Adverse Events Attributed to Study Treatment by the Investigator: Hypothyroidism: 135 (31.5%) Hyperthyroidism: 52 (12.1%) Adverse events of interest (with immune-related cause regardless of treatment attribution): Hypothyroidism: 152 (35.4%) Hyperthyroidism: 55 (12.8%) Thyroiditis: 12 (2.8%) (52)	In line with Keynote-011 (52)

Table S12. (continued)

Study (NCT number)	Frequency of thyroid screening	Definitions used to identify thyroid disorders
Keynote-048 (NCT02358031)	T3/FT3, FT4, and TSH: at screening, every other cycle, and safety follow-up visit (30 +-3 days following last dose) (53)	Data on adverse events and laboratory abnormalities were collected regularly throughout treatment and for 30 days thereafter (90 days for serious adverse events and events of interest) and graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (version 4.0) (53)

Incidence thyroid disorders	Definition used to identify immune-related adverse events
Pembrolizumab alone: Adverse events of any cause: Hypothyroidism: 55 (18%) Adverse events attributed by the physician to study treatment: Hypothyroidism: 39 (13%) Adverse events of interest (with immune-related cause regardless of treatment attribution): Hypothyroidism: 55 (18%) Hyperthyroidism: 8 (3%) Thyroiditis: 0 (0%) **Pembrolizumab with chemotherapy:** Adverse events of any cause: Hypothyroidism: 44 (16%) Adverse events attributed by the physician to study treatment Hypothyroidism: 36 (13%) Adverse events of interest (with immune-related cause regardless of treatment attribution): Hypothyroidism: 44 (16%) Hyperthyroidism: 12 (4%) Thyroiditis: 1 (<1%) (53)	In line with Keynote-011 (53)

REFERENCES SUPPLEMENTARY INFORMATION 2

1. Rittmeyer A, Barlesi F, Waterkamp D, Park K, Ciardiello F, von Pawel J, et al. Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial. Lancet. 2017;389(10066):255-65.
2. European Medicines Agency. Tecentriq: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/tecentriq-epar-public-assessment-report_en.pdf. Accessed 23/12/2020.
3. Center for Drug Evaluation and Research. Drug Approval Package: TECENTRIQ (atezolizumab) (761041Orig1s000). https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/761041Orig1s000MedR.pdf. Accessed 23/12/2020.
4. Fehrenbacher L, Spira A, Ballinger M, Kowanetz M, Vansteenkiste J, Mazieres J, et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. Lancet. 2016;387(10030):1837-46.
5. Center for Drug Evaluation and Research. Drug Approval Package: TECENTRIQ (atezolizumab) (761034Orig1s000) https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/761034Orig1s000MedR.pdf. Accessed 23/12/2020.
6. Balar AV, Galsky MD, Rosenberg JE, Powles T, Petrylak DP, Bellmunt J, et al. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. Lancet. 2017;389(10064):67-76.
7. Rosenberg JE, Hoffman-Censits J, Powles T, van der Heijden MS, Balar AV, Necchi A, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. Lancet. 2016;387(10031):1909-20.
8. Powles T, Durán I, van der Heijden MS, Loriot Y, Vogelzang NJ, De Giorgi U, et al. Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. Lancet. 2018;391(10122):748-57.
9. United States National Library of Medicine. A Study of Atezolizumab Compared With Chemotherapy in Participants With Locally Advanced or Metastatic Urothelial Bladder Cancer [IMvigor211]. https://clinicaltrials.gov/ct2/show/NCT02302807?term=NCT02302807&draw=2&rank=1. Accessed 23/12/2020.
10. 10. Socinski MA, Jotte RM, Cappuzzo F, Orlandi F, Stroyakovskiy D, Nogami N, et al. Atezolizumab for First-Line Treatment of Metastatic Nonsquamous NSCLC. N Engl J Med. 2018;378(24):2288-301.
11. Horn L, Mansfield AS, Szczęsna A, Havel L, Krzakowski M, Hochmair MJ, et al. First-Line Atezolizumab plus Chemotherapy in Extensive-Stage Small-Cell Lung Cancer. N Engl J Med. 2018;379(23):2220-9.
12. West H, McCleod M, Hussein M, Morabito A, Rittmeyer A, Conter HJ, et al. Atezolizumab in combination with carboplatin plus nab-paclitaxel chemotherapy compared with chemotherapy alone as first-line treatment for metastatic non-squamous non-small-cell lung cancer (IMpower130): a multicentre, randomised, open-label, phase 3 trial. Lancet Oncol. 2019;20(7):924-37.
13. Center for Drug Evaluation and Research. Drug Approval Package: Bavencio (avelumab) (761049Orig1s000). https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/761049Orig1s000MultidisciplineR.pdf. Accessed 28/12/2020.
14. Kaufman HL, Russell J, Hamid O, Bhatia S, Terheyden P, D'Angelo SP, et al. Avelumab in patients with chemotherapy-refractory metastatic Merkel cell carcinoma: a multicentre, single-group, open-label, phase 2 trial. Lancet Oncol. 2016;17(10):1374-85.
15. Motzer RJ, Penkov K, Haanen J, Rini B, Albiges L, Campbell MT, et al. Avelumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. N Engl J Med. 2019;380(12):1103-15.

16. Migden MR, Rischin D, Schmults CD, Guminski A, Hauschild A, Lewis KD, et al. PD-1 Blockade with Cemiplimab in Advanced Cutaneous Squamous-Cell Carcinoma. N Engl J Med. 2018;379(4):341-51.

17. Rischin D, Migden MR, Lim AM, Schmults CD, Khushalani NI, Hughes BGM, et al. Phase 2 study of cemiplimab in patients with metastatic cutaneous squamous cell carcinoma: primary analysis of fixed-dosing, long-term outcome of weight-based dosing. Journal for ImmunoTherapy of Cancer. 2020;8(1):e000775.

18. Migden MR, Khushalani NI, Chang ALS, Lewis KD, Schmults CD, Hernandez-Aya L, et al. Cemiplimab in locally advanced cutaneous squamous cell carcinoma: results from an open-label, phase 2, single-arm trial. The Lancet Oncology. 2020;21(2):294-305.

19. Antonia SJ, Villegas A, Daniel D, Vicente D, Murakami S, Hui R, et al. Durvalumab after Chemoradiotherapy in Stage III Non-Small-Cell Lung Cancer. N Engl J Med. 2017;377(20):1919-29.

20. AstraZeneca. Clinical Study Protocol: A Phase III, Randomized, Multicenter, Open-Label, Comparative Study to Determine the Efficacy of Durvalumab or Durvalumab and Tremelimumab in Combination With Platinum-Based Chemotherapy for the First-Line Treatment in Patients with Extensive Disease Small-Cell Lung Cancer (SCLC) (CASPIAN). https://s3.amazonaws.com/ctr-med-7111/D419QC00001/042de696-9399-4abc-8643-673f1237463a/7c52d3ce-065a-490b-b661-7a7c8c1c741f/d419qc00001-csp-v5_Redacted-v1.pdf. Accessed 28/12/2020.

21. Paz-Ares L, Dvorkin M, Chen Y, Reinmuth N, Hotta K, Trukhin D, et al. Durvalumab plus platinum-etoposide versus platinum-etoposide in first-line treatment of extensive-stage small-cell lung cancer (CASPIAN): a randomised, controlled, open-label, phase 3 trial. Lancet. 2019;394(10212):1929-39.

22. Robert C, Long GV, Brady B, Dutriaux C, Maio M, Mortier L, et al. Nivolumab in Previously Untreated Melanoma without BRAF Mutation. New England Journal of Medicine. 2014;372(4):320-30.

23. Weber JS, D'Angelo SP, Minor D, Hodi FS, Gutzmer R, Neyns B, et al. Nivolumab versus chemotherapy in patients with advanced melanoma who progressed after anti-CTLA-4 treatment (CheckMate 037): a randomised, controlled, open-label, phase 3 trial. Lancet Oncol. 2015;16(4):375-84.

24. Center for Drug Evaluation and Research. Drug Approval Package: OPDIVO (nivolumab) (125554Orig1s000). https://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/125554Orig1s000MedR.pdf. Accessed 5/1/2021.

25. Brahmer J, Reckamp KL, Baas P, Crinò L, Eberhardt WE, Poddubskaya E, et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. N Engl J Med. 2015;373(2):123-35.

26. European Medicines Agency. Opdivo-H-C-3985-II-0001 : EPAR - Assessment Report – Variation. https://www.ema.europa.eu/en/documents/variation-report/opdivo-h-c-3985-ii-0001-epar-assessment-report-variation_en.pdf. Accessed 5/1/2021.

27. Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, et al. Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. N Engl J Med. 2015;373(17):1627-39.

28. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. N Engl J Med. 2015;373(19):1803-13.

29. European Medicines Agency. Opdivo-H-C-3985-II-0008 : EPAR - Assessment Report – Variation. https://www.ema.europa.eu/en/documents/variation-report/opdivo-h-c-3985-ii-0008-epar-scientific-discussion-variation_en.pdf. Accessed 5/1/2021.

30. Larkin J, Chiarion-Sileni V, Gonzalez R, Grob JJ, Cowey CL, Lao CD, et al. Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. N Engl J Med. 2015;373(1):23-34.

31. United States National Library of Medicine. Study of Nivolumab in Patients With Classical Hodgkin's Lymphoma (Registrational) (CheckMate 205). https://clinicaltrials.gov/ProvidedDocs/38/NCT02181738/Prot_000.pdf. Accessed 5/1/2021.

32. Armand P, Engert A, Younes A, Fanale M, Santoro A, Zinzani PL, et al. Nivolumab for Relapsed/Refractory Classic Hodgkin Lymphoma After Failure of Autologous Hematopoietic Cell Transplantation: Extended Follow-Up of the Multicohort Single-Arm Phase II CheckMate 205 Trial. J Clin Oncol. 2018;36(14):1428-39.

33. Ferris RL, Blumenschein G, Jr., Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med. 2016;375(19):1856-67.

34. Sharma P, Retz M, Siefker-Radtke A, Baron A, Necchi A, Bedke J, et al. Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. Lancet Oncol. 2017;18(3):312-22.

35. European Medicines Agency. Opdivo-H-C-3985-II-0019 : EPAR - Assessment Report – Variation. https://www.ema.europa.eu/en/documents/variation-report/opdivo-h-c-3985-ii-0019-epar-assessment-report-variation_en.pdf. Accessed 5/1/2021.

36. Weber J, Mandala M, Del Vecchio M, Gogas HJ, Arance AM, Cowey CL, et al. Adjuvant Nivolumab versus Ipilimumab in Resected Stage III or IV Melanoma. N Engl J Med. 2017;377(19):1824-35.

37. Motzer RJ, Tannir NM, McDermott DF, Arén Frontera O, Melichar B, Choueiri TK, et al. Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma. N Engl J Med. 2018;378(14):1277-90.

38. European Medicines Agency. Opdivo-H-C-3985-WS-01278 : EPAR - Assessment Report – Variation. https://www.ema.europa.eu/en/documents/variation-report/opdivo-h-c-3985-ws-01278-epar-assessment-report-variation_en.pdf. Accessed 6/1/2021.

39. United States National Library of Medicine. A Study of Nivolumab and Ipilimumab Combined With Chemotherapy Compared to Chemotherapy Alone in First Line NSCLC (CheckMate 9LA). https://clinicaltrials.gov/ct2/show/study/NCT03215706. Accessed 5/1/2021.

40. Ribas A, Puzanov I, Dummer R, Schadendorf D, Hamid O, Robert C, et al. Pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory melanoma (KEYNOTE-002): a randomised, controlled, phase 2 trial. The Lancet Oncology. 2015;16(8):908-18.

41. Robert C, Schachter J, Long GV, Arance A, Grob JJ, Mortier L, et al. Pembrolizumab versus Ipilimumab in Advanced Melanoma. N Engl J Med. 2015;372(26):2521-32.

42. Herbst RS, Baas P, Kim DW, Felip E, Pérez-Gracia JL, Han JY, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet. 2016;387(10027):1540-50.

43. Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, et al. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. N Engl J Med. 2016;375(19):1823-33.

44. Chen R, Zinzani PL, Fanale MA, Armand P, Johnson NA, Brice P, et al. Phase II Study of the Efficacy and Safety of Pembrolizumab for Relapsed/Refractory Classic Hodgkin Lymphoma. J Clin Oncol. 2017;35(19):2125-32.

45. Bellmunt J, de Wit R, Vaughn DJ, Fradet Y, Lee JL, Fong L, et al. Pembrolizumab as Second-Line Therapy for Advanced Urothelial Carcinoma. N Engl J Med. 2017;376(11):1015-26.

46. United States National Library of Medicine. Study of Pembrolizumab (MK-3475) in Participants With Advanced Urothelial Cancer (MK-3475-052/KEYNOTE-052). https://clinicaltrials.gov/ct2/show/study/NCT02335424. Accessed 8/1/2021.

47. European Medicines Agency. Keytruda-H-C-2830-II-0023-G : EPAR - Assessment Report – Variation. https://www.ema.europa.eu/en/documents/variation-report/keytruda-h-c-2830-ii-0023-g-epar-assessment-report-variation_en.pdf. Accessed 8/1/2021.

48. Cohen EEW, Soulières D, Le Tourneau C, Dinis J, Licitra L, Ahn MJ, et al. Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study. Lancet. 2019;393(10167):156-67.

49. Gandhi L, Rodríguez-Abreu D, Gadgeel S, Esteban E, Felip E, De Angelis F, et al. Pembrolizumab plus Chemotherapy in Metastatic Non-Small-Cell Lung Cancer. N Engl J Med. 2018;378(22):2078-92.

50. Eggermont AMM, Blank CU, Mandala M, Long GV, Atkinson V, Dalle S, et al. Adjuvant Pembrolizumab versus Placebo in Resected Stage III Melanoma. N Engl J Med. 2018;378(19):1789-801.

51. Paz-Ares L, Luft A, Vicente D, Tafreshi A, Gümüş M, Mazières J, et al. Pembrolizumab plus Chemotherapy for Squamous Non-Small-Cell Lung Cancer. N Engl J Med. 2018;379(21):2040-51.

52. Rini BI, Plimack ER, Stus V, Gafanov R, Hawkins R, Nosov D, et al. Pembrolizumab plus Axitinib versus Sunitinib for Advanced Renal-Cell Carcinoma. N Engl J Med. 2019;380(12):1116-27.

53. Burtness B, Harrington KJ, Greil R, Soulières D, Tahara M, de Castro G, Jr., et al. Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-048): a randomised, open-label, phase 3 study. Lancet. 2019;394(10212):1915-28.

SUPPLEMENTARY INFORMATION 3. OBSERVATIONAL STUDIES

Table S1. Characteristics of the included observational studies.

Study	Treatment schedule	Types of malignancy	# patients included	Follow-up time	Time frame
de Filette et al. 2016 (1)	Pembrolizumab, 2 mg/kg Q3W	Advanced/unresectable melanoma	99	Median follow-up: 20.7 weeks (range: 1.1 – 72.6)	3 September 2014 – 4 January 2016
Delivanis et al. 2017 (2)	Pembrolizumab, 2 mg/kg Q3W	Not specified	93	Median follow-up: 8 months (range: 3 – 41)	April 2014 – January 2015
Maekura et al. 2017 (3)	Nivolumab	Non-small cell lung cancer	64	Median number of courses: 6 (range: 1-12)	December 2015 – May 2016
Morganstein et al. 2017 (4)	Nivolumab Pembrolizumab	Melanoma	46	Not specified	May 2013 – December 2014
Yamazaki et al. 2017 (5)	Nivolumab, 3 mg/kg Q2W	Non-small cell lung cancer	72	Median number of doses: 4 (range: 1-22)	January 2016 – December 2016
Kim et al. 2018 (6)	Nivolumab, 3 mg/kg Q2W Pembrolizumab, 2 mg/kg Q3W	Stage IV non-small cell lung cancer	58	Median follow-up: 89 days (IQR: 42-170)	January 2014 – December 2016
Kimbara et al. 2018 (7)	Nivolumab	Advanced solid tumors (including malignant melanoma, non-small cell lung cancer, and others)	168	Median follow-up: 272 days (range: 100-2529)	March 2009 – March 2016
Kobayashi et al. 2018 (8)	Nivolumab, 2 mg/kg Q3W Nivolumab, 3 mg/kg Q2W	Hodgkin lymphoma Non-small cell lung cancer Renal cell carcinoma Unresectable metastatic melanoma	66	Mean follow-up: 20 weeks (SD: 7)	2 November 2015 – 17 May 2017
Patel et al. 2018 (9)	Nivolumab Pembrolizumab	Any type of cancer, mostly melanoma, non-small cell lung cancer, head and neck cancer and renal cell carcinoma	139	Not specified	2010 – 2016

Table S1. (continued)

Study	Treatment schedule	Types of malignancy	# patients included	Follow-up time	Time frame
Scott et al. 2018 (11)	Nivolumab, Q2W Pembrolizumab, Q3W	Metastatic melanoma	103	Not specified	April 2014 – October 2015
Villa et al. 2018 (12)	Ipilimumab Nivolumab Pembrolizumab	Any malignant indication at an urban academic tertiary care institution	388	Not specified	2009 – 2016
Yano et al. 2018 (13)	Nivolumab, 3 mg/kg	Malignant melanoma	24	Mean administration duration: 32.1 weeks (SD: 31.4)	1 September 2014 – 30 September 2016
Campredon et al. 2019 (14)	Nivolumab, 3 mg/kg Q2W	Non-small cell lung cancer	105	Median follow-up: 9 months (95%CI: 7.5-10.3)	May 2015 – December 2016
Kassi et al. 2019 (15)	Atezolizumab, 1200 mg Ipilimumab, 3 mg/kg or 10 mg/kg Q3W Nivolumab, 3 mg/kg Q2W Nivolumab 1 mg/kg in combination with 3 mg/kg or 10 mg/kg ipilimumab Q3W Pembrolizumab 2 mg/kg Q3W	Melanoma	339	Median follow-up: 15.95 months (range: 5-57)	March 2014 – November 2018
Koyama et al. 2019 (16)	Nivolumab Pembrolizumab	Non-small cell lung cancer	132	Median follow-up: 7.6 months (range: 0.2-19.8)	December 2015 – June 2017

Table S1. (continued)

Study	Treatment schedule	Types of malignancy	# patients included	Follow-up time	Time frame
Peiro et al. 2019 (17)	Nivolumab, 3 mg/kg Q2W Nivolumab in combination with ipilimumab Nivolumab in combination with doxorubicin, vinblastine, and dacarbazine	Hodgkin lymphoma Non-small cell lung cancer Melanoma	73	Median follow-up: 364.5 days (IQR: 120-597.3)	During 2016
Pollack et al. 2019 (18)	Nivolumab Nivolumab in combination with ipilimumab Pembrolizumab Pembrolizumab in combination with ipilimumab	Metastatic melanoma	99	Not specified	1 November 2011 – 28 February 2017
Ramos-Levi et al. 2019 (10)	Nivolumab, 3 mg/kg Q2W	Advanced non-small cell lung cancer	40	Median follow-up: 7.6 ± 3.9 (1-15) months	February 2016 – April 2017
Sakakida et al. 2019 (19)	Nivolumab, 3 mg/kg Q2W Pembrolizumab, 2 mg/kg Q3W	Gastric cancer Head and neck cancer Hodgkin's lymphoma Malignant melanoma Non-small cell lung carcinoma Renal cell carcinoma Urothelial cancer	150	Median follow-up: 29 weeks (range: 4-203)	September 2014 – July 2018
Yamauchi et al. 2019 (20)	Nivolumab, 2 mg/kg Q3W Nivolumab, 3 mg/kg Q2W	Lung cancer Malignant melanoma Others	200	Median follow-up: 286 days (IQR: 106-479)	1 September 2014 – 31 August 2017

Table S1. (continued)

Study	Treatment schedule	Types of malignancy	# patients included	Follow-up time	Time frame
Al Mushref et al. 2020 (21)	Nivolumab Pembrolizumab	Metastatic melanoma	103	Not specified	2011 – 2017
Basak et al. 2020 (22)	Nivolumab, Q2W Pembrolizumab, Q3W	Metastatic melanoma Non-small-cell lung carcinoma Renal cell carcinoma	168	Median follow-up: 14.9 months (IQR: 9.2-18.4)	April 2016 – July 2017
Kotwal et al. 2020 (23)	Atezolizumab, 1200 mg Q3W Avelumab, 1200 mg Q3W	Lung, uroepithelial, Mercel cell, prostate, penis	91	Median follow-up: 10.1 months (IQR: 4.2 – 18.2)	1 June 2016 – 30 January 2018
Lima Ferreira et al. 2020 (24)	Nivolumab, dose according to current oncology protocols Pembrolizumab, dose according to current oncology protocols	Head and neck cancer Hodgkin's lymphoma Melanoma Lung cancer Urothelial carcinoma	161	Median follow-up: 49.4 weeks (26.5-75.8)	March 2014 – September 2019
Olsson-Brown et al. 2020 (25)	Nivolumab Pembrolizumab	Metastatic melanoma	90	Not specified	February 2016 – May 2017
Pollack et al. 2020 (26)	Atezolizumab Durvalumab Nivolumab Pembrolizumab	Lung, gastroenterological or genitourinary malignancies	188	Median follow-up: 12 months (IQR: 4-23)	January 2014 – December 2018

Table S1. (continued)

Study	Treatment schedule	Types of malignancy	# patients included	Follow-up time	Time frame
Presotto et al. 2020 (27)	Nivolumab, usually 3 mg/kg every 2 weeks Pembrolizumab, usually 2 mg/kg every 3 weeks	Bladder cancer Breast cancer Head-Neck cancer Kidney cancer Non-small cell lung cancer Melanoma	179	Follow-up: 2-33 months	15 July 2018
Sbardella et al. 2020 (28)	Nivolumab, 240 mg Pembrolizumab, 200 mg	Metastatic melanoma Non-small cell lung carcinoma Renal cell carcinoma	126	Mean follow-up: 10.2 months (SD: 10.9)	June 2015 – December 2018

Table S2. Description of the applied definitions to identify immune-related thyroid adverse events, frequency of screening of the thyroid levels and incidence of thyroid disorders in the observational studies.

Study	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders
de Filette et al. 2016	At baseline and before each pembrolizumab administration	Subclinical thyrotoxicosis: TSH <0.27 mIU/L and fT4 11.6-22.0 pmol/L and/or fT3 4.0-6.8 pmol/L Thyrotoxicosis: TSH <0.27 mIU/L and fT4 >22.0 pmol/L and/or fT3 >4.0 pmol/L Subclinical hypothyroidism: TSH >4.2 mIU/L and fT4 11.6-22.0 pmol/L and/or fT3 4.0-6.8 pmol/L Hypothyroidism: TSH >4.2 mIU/L and fT4 <11.6 pmol/L	Subclinical thyrotoxicosis: 1 (1.0%) Thyrotoxicosis: 11 (11.1%) Subclinical hypothyroidism 3 (3.0%) Hypothyroidism: 12 (15.2%) Total: 17 (17.2%)
Delivanis et al. 2017	At baseline and every 3 weeks while on treatment	Thyroiditis: TSH <0.3 mIU/L level with normal or elevated T4 or triiodothyronine that spontaneously resolved or progressed to overt hypothyroidism Subclinical hypothyroidism: TSH >4.2 but ≤ 10 mIU/L with normal serum FT4 (11.6-21.9 pmol/L) Overt hypothyroidism: TSH >10 with or without low FT4 Recurrent hypothyroidism: patients with a history of hypothyroidism and on stable thyroid hormone replacement who, following initiation of pembrolizumab, developed an acute rise in TSH requiring an increased in their levothyroxine dose by >50%.	Thyroiditis: 7 (7.5%) Subclinical hypothyroidism: 1 (1.1%) Overt hypothyroidism: 2 (2.2%) Recurrent hypothyroidism: 3 (3.2%) Total: 13 (14.0%)
Maekura et al. 2017	Not specified	Hypothyroidism: TSH >4.2 mIU/L with FT4 <12.87 pmol/L	Hypothyroidism: 5 (7.8%)

Table S2. (continued)

Study	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders
Morganstein et al. 2017	At baseline and then every 2-4 weeks	Subclinical hyperthyroidism: TSH <0.49 mIU/L with FT4 9.1–23.8 pmol/L Hyperthyroidism: FT4 >23.8 pmol/L Subclinical hypothyroidism: TSH >4.67 mIU/L with FT4 9.1–23.8 pmol/L Hypothyroidism: FT4 <9.1 pmol/L	Subclinical hyperthyroidism: 6 (13.0%) Subclinical hypothyroidism: 6 (13.0%) Hypothyroidism: 6 (13.0%) Total: 18 (39.1%)
Yamazaki et al. 2017	Not specified	Subclinical hyperthyroidism: decreased TSH with fT4 within the reference range Subclinical hypothyroidism: elevated TSH with fT4 within the reference range	(Subclinical) hyperthyroidism: 11 (15.3%) (Subclinical) hypothyroidism: 3 (4.2%) Total: 14 (19.4%)
Kim et al. 2018	Every 4-8 weeks	Subclinical thyrotoxicosis: TSH <0.30 mIU/L with fT4 10.16-23.94 pmol/L for male, 8.23-22.13 pmol/L for female Overt thyrotoxicosis: TSH <0.30 mIU/L with fT4 >23.94 pmol/L for male, >22.13 pmol/L for female Subclinical hypothyroidism: TSH >6.00 mIU/L for male, >6.50 mIU/L for female with fT4 10.16-23.94 pmol/L for male, 8.23-22.13 pmol/L for female Overt hypothyroidism: TSH >6.00 mIU/L for male, >6.50 mIU/L for female with fT4 <10.16 pmol/L for male, <8.23 pmol/L for female	Subclinical thyrotoxicosis: 3 (5.2%) Overt thyrotoxicosis: 5 (8.6%) Subclinical hypothyroidism: 6 (10.3%) Overt hypothyroidism: 5 (8.6%) Total: 19 (32.8%)

Table S2. (continued)

Study	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders
Kimbara et al. 2018	At baseline, every 1 or 2 months	Thyrotoxicosis: TSH <0.35 mIU/L with FT4 >19.05 pmol/L and/or FT3 >5.70 pmol/L Hypothyroidism: TSH >10 mIU/L with FT4 <9.01 pmol/L	Thyrotoxicosis: 20 (11.9%) Hypothyroidism: 17 (10.1%) Total: 23 (13.7%)
Kobayashi et al. 2018	At baseline and every 6 weeks after the first administration of nivolumab for 24 weeks.	Destructive thyroiditis: suppressed TSH level (<0.35 mIU/L) with an elevated FT3 (>5.70 pmol/L) and/or FT4 (>19.1 pmol/L) and no TSH receptor antibodies	Destructive thyroiditis: 4 (6.1%)
Patel et al. 2018	At baseline, every 2 or 3 weeks while on treatment	Subclinical hyperthyroidism: TSH <0.27 mIU/L with FT4 11.97-21.88 pmol/L Subclinical hypothyroidism: TSH >4.2 mIU/L with FT4 11.97-21.88 pmol/L Primary hypothyroidism: TSH >4.2 mIU/L with FT4 <11.97 pmol/L Primary overt hyperthyroidism versus thyroiditis-like syndrome: patients who presented initially with hyperthyroidism and progressed either to a euthyroid state or primary hypothyroidism with or without methimazole	Subclinical hyperthyroidism: 5 (3.6%) Subclinical hypothyroidism: 14 (10.1%) Primary hypothyroidism: 16 (11.5%) Primary overt hyperthyroidism versus thyroiditis-like syndrome: 4 (2.9%)

Table S2. (continued)

Study	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders
Scott et al. 2018	Prior to commencement of therapy and prior to each subsequent cycle for at least the first 12 weeks of treatment	All patients with clinical or biochemical endocrine abnormalities were referred to a single endocrinologist. Baseline clinicopathologic data, treatment and outcome data were examined.	Hyperthyroidism: 7 (6.8%) Subclinical hypothyroidism: 2 (1.9%) Total: 9 (8.7%)
Villa et al. 2018	Not specified	Subclinical hyperthyroidism: TSH <0.3 mIU/mL and [fT4 10.29– 20.59 pmol/L or FTI >135.15 nmol/L or TT3 >2.85nmol/L or fT3 >2.49pmol/L] Overt hyperthyroidism: TSH <0.3 mIU/mL and fT4 >20.59 pmol/L Subclinical hypothyroidism: TSH >4.7 mIU/mL and [fT4 3.86–20.59 pmol/L or FTI <57.92nmol/L or TT3 <1.31nmol/L or fT3 <2.49pmol/L] Overt hypothyroidism: TSH >4.7 mIU/mL and [fT4 <3.86pmol/L or FTI <57.92nmol/L or TT3 <1.31nmol/L or fT3 <2.49pmol/L] Upon review, patients who received treatment were reclassified as clinical hypothyroidism or hyperthyroidism instead of subclinical hypothyroidism or hyperthyroidism regardless of lab values.	Subclinical hyperthyroidism: 3 (0.8%) Overt hyperthyroidism: 9 (2.3%) Subclinical hypothyroidism: 21 (5.4%) Overt hypothyroidism: 10 (2.6%)
Yano et al. 2018	At baseline, every 3 weeks	Thyrotoxicosis: TSH <0.27 mIU/L and fT4 >23.17 pmol/L Subclinical hypothyroidism: TSH >4.2 mIU/L and fT4 12.87-23.17 pmol/L Hypothyroidism: TSH >4.2 mIU/L and fT4 <12.87 pmol/L	Thyrotoxicosis: 3 (12.5%) of which 2 later developed overt hypothyroidism and 1 subclinical hypothyroidism Subclinical hypothyroidism: 3 (12.5%) Overt hypothyroidism: 1 (4.2%) Total: 7 (29.2%)

Table S2. (continued)

Study	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders
Campredon et al. 2019	At baseline, before each infusion	Isolated thyrotoxicosis: TSH <0.1 mIU/mL, transient or permanent (persistent until the end of the study), with no increase of TSH level during monitoring Biphasic thyroiditis: transient thyrotoxicosis, with a decrease of TSH <0.1 mIU/ml, followed by hypothyroidism with an increase of TSH level above the patient's usual laboratory reference range Isolated hypothyroidism: increase of TSH level, above the patient's usual laboratory reference range (i.e. greater than 4 mIU/mL) and not preceeded by a thyrotoxic phase, transient or permanent at the end of the study. Hypothyroidism is transient when the increase of TSH level is followed by a spontaneous return to the normal level (or after stopping treatment with levothyroxine). It is permanent when the increase of TSH lasts more than 3 months or requires a longterm treatment with levothyroxine.	Isolated thyrotoxicosis 8 (7.6%) of which 7 transient Biphasic thyroiditis: 5 (4.8%) of which 1 transient Isolated hypothyroidism: 2 (1.9%) of which 1 transient Total: 15 (14.3%)
Kassi et al. 2019	Before initiation of the treatment and before every other cycle	Primary thyroid dysfunction: low levels of fT4 and/or fT3 with increased levels of TSH	Primary thyroid dysfunction: 19 (5.6%)
Koyama et al. 2019	Before treatment and monthly after the first administration	Immune-related thyroid dysfunction: two or more successive abnormal levels of TSH during anti-PD-1 treatment	Immune-related thyroid dysfunction: 19 (14.4%)

Table S2. (continued)

Study	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders
Peiro et al. 2019	At baseline and usually before each cycle or as it was indicated by clinical trial protocols	Subclinical thyrotoxicosis: TSH <0.48 mIU/L with FT4 9.7–30.9 pmol/L and/or FT3 3.1–6.8 pmol/L Thyrotoxicosis: TSH <0.48 mIU/L with FT4 >30.9 pmol/L and or FT3 >6.8 pmol/L Subclinical hypothyroidism: TSH >4.36 mIU/L with FT4 9.7–30.9 pmol/L Hypothyroidism: TSH >4.36 mIU/L with FT4 <9.7 pmol/L	Subclinical thyrotoxicosis: 2 (2.7%) Thyrotoxicosis: 5 (6.8%) Thyroiditis (hyperthyroid patients that became hypothyroid later): 5 (6.8%) Subclinical hypothyroidism: 8 (11.0%) Hypothyroidism: 2 (2.7%) Total: 17 (23.3%)
Pollack et al. 2019	Not specified	Any thyroid dysfunction: one or more consecutive abnormal TSH values (reference range: 0.5-4.5 mIU/L), irrespective of FT4 and FT3 levels, observed up to 120 days following the index date Overt hyperthyroidism: TSH <0.5 with FT4 >20.0 pmol/L and/or FT3 >6.5 pmol/L Overt hypothyroidism: TSH >4.5 mIU/L with FT4 <10.0 pmol/L and/or FT3 <3.5 pmol/L or TSH >10 mIU/L	Any thyroid dysfunction: 42 (42.4%) including 20 (20.2%) with subclinical thyroid dysfunction Overt hyperthyroidism: 9 (9.1%) of which 8 progressed to overt hypothyroidism, consistent with thyroiditis Overt hypothyroidism: 13 (13.1%)
Ramos-Levi et al. 2019	At baseline and monthly during subsequent follow-up visits	Immune-related thyroiditis in the hyperthyroid phase: suppressed TSH level (<0.5 mIU/L) with an elevated FT4 (>23.2 pmol/L) Immune-related primary hypothyroidism: TSH level ≥5 mIU/l alone with or without a low FT4.	Hyperthyroidism: 3 (7.5%) Hypothyroidism: 6 (15%) Total: 9 (22.5%)

Table S2. (continued)

Study	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders
Sakakida et al. 2019	Not specified	Thyroid dysfunction: at least two consecutive abnormal TSH level measurements (reference range: 0.5-5.0 mIU/L) Subclinical hypothyroidism: TSH >5.0 mIU/L with FT4 11.58-21.88 pmol/L Overt hypothyroidism: TSH >5.0 mIU/L with FT4 <11.58 pmol/L	Thyroid dysfunction: 25 (16.7%) of which 21 newly developed thyroid dysfunction of which 9 underwent an initial period of transient hyperthyroidism Subclinical hypothyroidism: 9 (6%) Overt hypothyroidism: 12 (8%)
Yamauchi et al. 2019	At treatment initiation, every 3 weeks in patients with malignant melanoma and every 4 weeks in patients with lung cancer and other malignancies	Thyroid immune-related adverse events: abnormal values of thyroid levels, excluding patients for which these abnormal values were not a result of nivolumab therapy according to the following criteria: 1) data of thyroid function tests were unavailable; 2) serum TSH levels were normalized during the follow-up period; 3) total thyroidectomy was previously performed; 4) serum TSH levels did not exceed 10 mIU/mL in patients with mild subclinical hypothyroidism (defined as TSH <10 mIU/mL); 5) serum fT4 levels did not fall below the reference range in patients with moderate subclinical hypothyroidism (defined as TSH ≥10 mIU/mL); and 6) low TSH values were only observed during glucocorticoid therapy. Overt thyroid immune-related adverse events: neither FT4 and TSH levels were normal (reference range: FT4 11.33-20.85 pmol/L, TSH: 0.5-5.0 mIU/L) Subclinical thyroid immune-related adverse events: FT4 or TSH levels were abnormal	Thyroid immune-related adverse events: 67 (33.5%) Overt thyroid immune-related adverse events: 27 (13.5%) of which for 17 patients transient thyrotoxicosis was observed which was followed by hypothyroidism in 11 patients Subclinical immune-related adverse events: 40 (20%)
Al Mushref et al. 2020	Not specified	Thyroid dysfunction: TSH >5 mIU/L or TSH <0.5 mIU/L. Excluding: thyroid dysfunction prior to immunotherapy, evidence of pituitary dysfunction (likely checkpoint inhibitor hypophysitis), concurrent neck radiation or use of amiodarone, and less than two documented thyroid function abnormalities following checkpoint inhibitors	Thyroid dysfunction: 8 (7.8%)

Table S2. (continued)

Study	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders
Basak et al. 2020	At baseline, before each infusion	Subclinical hyperthyroidism: TSH <0.4 mIU/L with fT4 11-25 pmol/L Overt hyperthyroidism: TSH <0.4 mIU/L with fT4 >25 pmol/L Subclinical hypothyroidism: TSH >4.3 mIU/L with fT4 11-25 pmol/L Overt hypothyroidism: TSH >4.3 mIU/L with fT4 <11 pmol/L	Subclinical hyperthyroidism: 12 (7.1%) Overt hyperthyroidism: 14 (8.3%) Subclinical hypothyroidism: 22 (13.1%) Overt hypothyroidism: 6 (3.6%) Total: 54 (32.1%)
Kotwal et al. 2020	Not specified	Thyroid immune-related adverse events: two or more abnormal thyroid function tests after starting a PD-L1 inhibitor, and in the absence of other causes and further characterized as: Thyrotoxicosis, subclinical: TSH ≤0.2 mIU/L and free T4 11.6-21.9 pmol/L Thyrotoxicosis, overt: TSH ≤0.2 mIU/L and free T4 ≥23.2 pmol/L New onset hypothyroidism, subclinical: TSH ≥4.3 mIU/L and free T4 11.6-21.9 pmol/L New onset hypothyroidism, overt: TSH ≥4.3 mIU/L and free T4 ≤10.30 pmol/L Worsening of preexisting hypothyroidism: patients with a history of hypothyroidism and on stable thyroid hormone replacement who, after initiation of the PD-L1 inhibitor, developed an acute rise in TSH requiring an increase in their levothyroxine dose by >50%	Thyrotoxicosis, subclinical: 4 (4.4%) Thyrotoxicosis, overt: 1 (1.1%) New onset hypothyroidism, subclinical: 10 (11.0%) New onset hypothyroidism, overt: 4 (4.4%) Worsening of preexisting hypothyroidism: 4 (4.4%) Total: 23 (25.3%)

Table S2. (continued)

Study	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders
Lima Ferreira et al. 2020	Every treatment visit or commonly at most every four weeks	Subclinical thyrotoxicosis: TSH <0.10 mIU/L with FT4 11.97-21.88 pmol/L and/or FT3 3.07-6.76 pmol/L Overt thyrotoxicosis: TSH <0.10 mIU/L with FT4 >21.88 pmol/L and/or FT3 >6.76 pmol/L Subclinical hypothyroidism: TSH >4.2 mIU/L with FT4 11.97-21.88 pmol/L and/or FT3 3.07-6.76 pmol/L Overt hypothyroidism: TSH >4.2 mIU/L with FT4 <11.97 pmol/L and/or FT3 <3.07 pmol/L Biphasic thyroiditis: transient thyrotoxicosis followed by hypothyroidism	Thyrotoxicosis: 4 (2.6%) Subclinical hypothyroidism: 4 (2.6%) Overt hypothyroidism: 8 (5.2%) Biphasic thyroiditis: 4 (2.6%)
Olsson-Brown et al. 2020	At baseline and recurrent on a 3-weekly basis	Thyroid dysfunction: presence of deranged thyroid function denoted by a T4 level outside the standard reference range (11.5-22.7 pmol/L). All patients found to have a single abnormal T4 level were evaluated to determine if the changes were transient or led to established thyroid dysfunction.	Thyroid dysfunction: 13 (14.4%) Hyperthyroidism followed by hypothyroidism: 9 (10%) De novo hypothyroidism: 4 (4.4%)

Table S2. (continued)

Study	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders
Pollack et al. 2020	Not measured at set intervals. In the initial study years, they were measured per physician discretion; due to increasing awareness of the possibility of thyroid dysfunction due to therapy, frequent tests were taken. In more recent years, with increasing awareness, TFTs were taken more frequently, nearly prior to each treatment cycle	Any thyroid dysfunction: TSH value above or below the laboratory-specific range, irrespective of FT4 and FT3 levels during follow-up. Any subclinical thyroid dysfunction: TSH above the laboratory-specific reference range but <10 mU/L, or a TSH value below the laboratory-specific reference range without a corresponding FT4 or FT3 abnormality Overt thyrotoxicosis was defined as a suppressed TSH with elevated FT4 and/or FT3 levels Overt hypothyroidism: either elevated TSH with reduced FT4 and/or FT3 or TSH ≥10 mU/L Thyroiditis: as subclinical or overt thyrotoxicosis followed by progression to overt hypothyroidism	Any thyroid dysfunction: 72 (38.9%) Any subclinical thyroid dysfunction: 31 (16.8%) Overt thyrotoxicosis: 17 (9.2%) Overt hypothyroidism: 32 (17.3%) Thyroiditis: 12 (6.5%)

Table S2. (continued)

Study	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders
Presotto et al. 2020	As part of clinical practice	Subclinical thyrotoxicosis: TSH <0.3 mIU/L with fT4 10–19 pmol/L and fT3 3.1–6.8 pmol/L Overt thyrotoxicosis: TSH <0.3 mIU/ with fT4 >19 pmol/L and fT3 >6.8 pmol/L Subclinical hypothyroidism: TSH >4.5 mIU/L with fT4 10–19 pmol/L and fT3 3.1–6.8 pmol/L Overt hypothyroidism: TSH >4.5 mIU/L with fT4 <10 pmol/L and fT3 <3.1 pmol/L	Subclinical hyperthyroidism: 11 (6.1%) Hyperthyroidism: 2 (1.1%) Subclinical hypothyroidism: 20 (11.2%) Overt hypothyroidism: 4 (2.2%) Patients with biphasic pattern: Subclinical hyperthyroidism: 9 (5.0%) Hyperthyroidism: 7 (3.9%) Subclinical hypothyroidism: 12 (6.7%) Overt hypothyroidism: 3 (1.7%) Total: 53 (29.6%)

Table S2. (continued)

Study	Frequency of thyroid screening	Definitions used to identify thyroid disorders	Incidence thyroid disorders
Sbardella et al. 2020	At baseline and monthly for the duration of treatment	Subclinical hyperthyroidism: TSH <0.4 mIU/L with normal serum FT3 (3.07-6.76 pmol/L) and normal FT4 (12.0-21.9 pmol/L) Overt hyperthyroidism: TSH <0.4 mIU/L with high serum FT3 (>6.76 pmol/L) and/or FT4 (>21.9 pmol/L) Subclinical hypothyroidism: TSH >10 mIU/L with normal serum FT3 (3.07-6.76 pmol/L) and normal FT4 (12.0-21.9 pmol/L) Overt hypothyroidism: TSH >10 mIU/L with low serum FT3 (<3.07 pmol/L) and/or FT4 levels (<12.0 pmol/L)	Subclinical hyperthyroidism: 6 (4.8%) Overt hyperthyroidism: 4 (3.2%) Subclinical hypothyroidism: 15 (11.9%) Overt hypothyroidism: 4 (3.2%) Total: 29 (23%)

REFERENCES SUPPLEMENTARY INFORMATION 3

1. de Filette J, Jansen Y, Schreuer M, Everaert H, Velkeniers B, Neyns B, et al. Incidence of Thyroid-Related Adverse Events in Melanoma Patients Treated With Pembrolizumab. J Clin Endocrinol Metab. 2016;101(11):4431-9.
2. Delivanis DA, Gustafson MP, Bornschlegl S, Merten MM, Kottschade L, Withers S, et al. Pembrolizumab-Induced Thyroiditis: Comprehensive Clinical Review and Insights Into Underlying Involved Mechanisms. J Clin Endocrinol Metab. 2017;102(8):2770-80.
3. Maekura T, Naito M, Tahara M, Ikegami N, Kimura Y, Sonobe S, et al. Predictive Factors of Nivolumab-induced Hypothyroidism in Patients with Non-small Cell Lung Cancer. In Vivo. 2017;31(5):1035-9.
4. Morganstein DL, Lai Z, Spain L, Diem S, Levine D, Mace C, et al. Thyroid abnormalities following the use of cytotoxic T-lymphocyte antigen-4 and programmed death receptor protein-1 inhibitors in the treatment of melanoma. Clin Endocrinol (Oxf). 2017;86(4):614-20.
5. Yamazaki H, Iwasaki H, Yamashita T, Yoshida T, Suganuma N, Yamanaka T, et al. Potential Risk Factors for Nivolumab-induced Thyroid Dysfunction. In vivo (Athens, Greece). 2017;31(6):1225-8.
6. Kim HI, Kim M, Lee SH, Park SY, Kim YN, Kim H, et al. Development of thyroid dysfunction is associated with clinical response to PD-1 blockade treatment in patients with advanced non-small cell lung cancer. Oncoimmunology. 2017;7(1):e1375642.
7. Kimbara S, Fujiwara Y, Iwama S, Ohashi K, Kuchiba A, Arima H, et al. Association of antithyroglobulin antibodies with the development of thyroid dysfunction induced by nivolumab. Cancer Sci. 2018;109(11):3583-90.
8. Kobayashi T, Iwama S, Yasuda Y, Okada N, Tsunekawa T, Onoue T, et al. Patients With Antithyroid Antibodies Are Prone To Develop Destructive Thyroiditis by Nivolumab: A Prospective Study. J Endocr Soc. 2018;2(3):241-51.
9. Patel NS, Oury A, Daniels GA, Bazhenova L, Patel SP. Incidence of Thyroid Function Test Abnormalities in Patients Receiving Immune-Checkpoint Inhibitors for Cancer Treatment. Oncologist. 2018;23(10):1236-41.
10. Ramos-Levi AM, Rogado J, Sanchez-Torres JM, Colomer R, Marazuela M. Nivolumab-induced thyroid dysfunction in patients with lung cancer. Endocrinologia, diabetes y nutricion. 2019;66(1):26-34.
11. Scott ES, Long GV, Guminski A, Clifton-Bligh RJ, Menzies AM, Tsang VH. The spectrum, incidence, kinetics and management of endocrinopathies with immune checkpoint inhibitors for metastatic melanoma. Eur J Endocrinol. 2018;178(2):173-80.
12. Villa NM, Farahmand A, Du L, Yeh MW, Smooke-Praw S, Ribas A, et al. Endocrinopathies with use of cancer immunotherapies. Clinical endocrinology. 2018;88(2):327-32.
13. Yano S, Ashida K, Nagata H, Ohe K, Wada N, Takeichi Y, et al. Nivolumab-induced thyroid dysfunction lacking antithyroid antibody is frequently evoked in Japanese patients with malignant melanoma. BMC Endocr Disord. 2018;18(1):36.
14. Campredon P, Mouly C, Lusque A, Bigay-Game L, Bousquet E, Mazières J, et al. Incidence of thyroid dysfunctions during treatment with nivolumab for non-small cell lung cancer: Retrospective study of 105 patients. Presse medicale (Paris, France : 1983). 2019;48(4):e199-e207.
15. Kassi E, Angelousi A, Asonitis N, Diamantopoulos P, Anastasopoulou A, Papaxoinis G, et al. Endocrine-related adverse events associated with immune-checkpoint inhibitors in patients with melanoma. Cancer medicine. 2019;8(15):6585-94.
16. Koyama J, Horiike A, Yoshizawa T, Dotsu Y, Ariyasu R, Saiki M, et al. Correlation between thyroid transcription factor-1 expression, immune-related thyroid dysfunction, and efficacy of anti-programmed cell death protein-1 treatment in non-small cell lung cancer. Journal of thoracic disease. 2019;11(5):1919-28.

17. Peiró I, Palmero R, Iglesias P, Díez JJ, Simó-Servat A, Marín JA, et al. Thyroid dysfunction induced by nivolumab: searching for disease patterns and outcomes. Endocrine. 2019;64(3):605-13.

18. Pollack RM, Kagan M, Lotem M, Dresner-Pollak R. BASELINE TSH LEVEL IS ASSOCIATED WITH RISK OF ANTI-PD-1-INDUCED THYROID DYSFUNCTION. Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists. 2019;25(8):824-9.

19. Sakakida T, Ishikawa T, Uchino J, Chihara Y, Komori S, Asai J, et al. Clinical features of immune-related thyroid dysfunction and its association with outcomes in patients with advanced malignancies treated by PD-1 blockade. Oncology letters. 2019;18(2):2140-7.

20. Yamauchi I, Yasoda A, Matsumoto S, Sakamori Y, Kim YH, Nomura M, et al. Incidence, features, and prognosis of immune-related adverse events involving the thyroid gland induced by nivolumab. PloS one. 2019;14(5):e0216954.

21. Al Mushref M, Guido PA, Collichio FA, Moore DT, Clemmons DR. THYROID DYSFUNCTION, RECOVERY, AND PROGNOSIS IN MELANOMA PATIENTS TREATED WITH IMMUNE CHECKPOINT INHIBITORS: A RETROSPECTIVE REVIEW. Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists. 2020;26(1):36-42.

22. Basak EA, van der Meer JWM, Hurkmans DP, Schreurs MWJ, Oomen-de Hoop E, van der Veldt AAM, et al. Overt Thyroid Dysfunction and Anti-Thyroid Antibodies Predict Response to Anti-PD-1 Immunotherapy in Cancer Patients. Thyroid. 2020;30(7):966-73.

23. Kotwal A, Kottschade L, Ryder M. PD-L1 Inhibitor-Induced Thyroiditis Is Associated with Better Overall Survival in Cancer Patients. Thyroid. 2020;30(2):177-84.

24. Lima Ferreira J, Costa C, Marques B, Castro S, Victor M, Oliveira J, et al. Improved survival in patients with thyroid function test abnormalities secondary to immune-checkpoint inhibitors. Cancer immunology, immunotherapy : CII. 2020.

25. Olsson-Brown A, Lord R, Sacco J, Wagg J, Coles M, Pirmohamed M. Two distinct clinical patterns of checkpoint inhibitor-induced thyroid dysfunction. Endocrine connections. 2020;9(4):318-25.

26. Pollack R, Ashash A, Cahn A, Rottenberg Y, Stern H, Dresner-Pollak R. Immune Checkpoint Inhibitor-induced Thyroid Dysfunction Is Associated with Higher Body Mass Index. The Journal of clinical endocrinology and metabolism. 2020;105(10).

27. Presotto EM, Rastrelli G, Desideri I, Scotti V, Gunnella S, Pimpinelli N, et al. Endocrine toxicity in cancer patients treated with nivolumab or pembrolizumab: results of a large multicentre study. Journal of Endocrinological Investigation. 2020;43(3):337-45.

28. Sbardella E, Tenuta M, Sirgiovanni G, Gianfrilli D, Pozza C, Venneri MA, et al. Thyroid disorders in programmed death 1 inhibitor-treated patients: Is previous therapy with tyrosine kinase inhibitors a predisposing factor? Clinical Endocrinology. 2020;92(3):258-65.

4.2

ADVERSE EVENTS RELATED TO BIOLOGICALS USED FOR PATIENTS WITH MULTIPLE SCLEROSIS: A COMPARISON BETWEEN INFORMATION ORIGINATING FROM REGULATORS AND INFORMATION ORIGINATING FROM THE SCIENTIFIC COMMUNITY

Lotte A. Minnema, Thijs J. Giezen, Toine C.G. Egberts,
Hubert G.M. Leufkens, Helga Gardarsdottir

Eur J Neurol. 2020 Jul;27(7):1250-1256

ABSTRACT

Background: Clinical decision making is facilitated by healthcare professionals' and patients' adequate knowledge of the adverse events. This is especially important for biologicals used for treating multiple sclerosis (MS). So far, little is known about whether different information sources report adverse events consistently.

Methods: Biologicals authorized by the European Medicines Agency for the treatment of MS were included in this study. Information on adverse events derived from phase 3 clinical trials from European Public Assessment Reports (EPARs) and from scientific publications was compared.

Results: In the study, eight biologicals used for the treatment of MS were included for which the EPAR and/or scientific publication reported a total of 707 adverse events. Approximately one-third of the adverse events was reported in both the EPAR and scientific publication, one-third was only reported in the EPAR and one-third only in the scientific publication. Serious adverse events and adverse events that regulators classified as 'important identified risk' were significantly more often reported in both sources compared to adverse events not classified as such (respectively, 38% vs. 30% and 49% vs. 30%). Adverse events only reported in the EPAR or in the scientific publication were, in general, not described in the benefit–risk section or abstract, which were considered to be the most important sections of the documents.

Conclusion: This study showed that there is substantial discordance in the reporting of adverse events on the same phase 3 trials between EPARs and scientific publications. To support optimal clinical decision making, both documents should be considered.

INTRODUCTION

Regulators have approved several biologicals to treat patients with relapsing and progressive multiple sclerosis (MS) during the last decade. Although these biologicals improve clinical symptoms and reduce relapse rates and disease progression, serious adverse events (SAEs) can occur. The detection of the adverse events (AEs) of these drugs may be complicated as these AEs can mimic the clinical expression of MS. For example, the early symptoms of encephalitis associated with the use of daclizumab include aphasia, confusion and disorientation, which are symptoms similar to those associated with a serious MS relapse (1). Encephalitis was therefore first interpreted as a worsening of the disease and as lack of efficacy of the drug instead of a SAE (1).

Healthcare professionals and patients can use different sources of information to obtain knowledge about the efficacy and safety profile of a drug in order to guide clinical decision making. At the time of approval, knowledge about the efficacy and safety profile is mainly based on the findings of the phase 3 randomized clinical trials that supported marketing approval. The results of these clinical trials are (publicly) available in various information sources. One of these information sources is peer-reviewed scientific publications where investigators report the results of the clinical trials. These scientific publications were an important source of evidence for the development of the European Clinical Guideline on the pharmacological treatment of people with MS (2). Another source is the publicly available European Public Assessment Report (EPAR). The European Medicines Agency (EMA), which is the regulatory authority in Europe responsible for evaluating marketing approval applications, publishes the EPAR; it provides an overview of the assessment procedure, including an assessment of the conducted clinical trials (3).

Although these two information sources reflect information obtained from the same clinical trials, the choice of the clinical findings that are extracted from these trials and the attention given to those clinical findings can differ. However, one might expect that the most important information generated from the clinical trials is reported in both documents. Several studies have assessed synergies between the reporting of efficacy and safety information from clinical trials by regulatory authorities and in scientific publications (4-9). These show that there are large differences in reporting between these two types of information sources. For example, de Vries et al. showed that, for antidepressants, 79% of the scientific publications provided incomplete information on SAEs compared to data obtained from the US Food and Drug Administration, and 63% did not mention SAEs at all (5). Another study on insomnia medication showed that scientific publications from studies identified in the EPAR reported reliably on the primary end-points but less reliably or not at all on the safety of the drug (6).

Since SAEs have occurred in clinical practice for biologicals used for MS, clinicians should have a comprehensive view of the safety profile to support clinical decision making. Therefore, the aim of this study was to provide, for biologicals used in MS, an analysis on which AEs from clinical trials are reported in the EPARs and the corresponding scientific publications, and whether these differ.

METHODS

Study drugs and information sources

In this study, biologicals that were or had been approved by the EMA for the treatment of MS (as of 31 December 2018) were included. The EPARs were retrieved from the EMA website (www.ema.europa.eu). The corresponding scientific publications of the phase 3 randomized clinical trials that supported approval of the product were identified using PubMed and the webpage clinicaltrials.gov. The full text of the scientific publication was obtained from the scientific journal concerned. Whether the scientific publications corresponded with the clinical trials described in the EPARs was verified by comparing the identifiers used in the EPARs and scientific publications (e.g. the clinicaltrials.gov identifier), the study design and the number of patients included. Furthermore, a cross-check with the Cochrane review on immunomodulators and immunosuppressants for relapsing–remitting MS was performed (10).

For each product, information on the year of approval, number of clinical trials supporting the approval of the product, and mechanism of action from the EPAR was retrieved.

Adverse events

For both information sources, the reported AEs for each product were compared. For the EPAR, the analysis was limited to the sections reporting on the safety information from the clinical trials and the benefit–risk discussion, whereas for the scientific publications all sections (including appendices, if applicable) were taken into account.

The AEs reported for the product were identified and characterized using the Medical Dictionary for Regulatory Affairs (MedDRA®) (11). MedDRA® is a validated standardized terminology used to facilitate the exchange of information on AEs, and it is used, amongst other things, in the communication of information from clinical trials between industry and regulators. MedDRA® is hierarchically structured. The lowest, and most specific, level reflects how an AE is reported in practice. Each of these lower level terms is linked to one preferred term. Multiple lower level terms can fall within one preferred term, as they may include synonyms or different word forms for the same expression. For example, the lower level terms 'multiple sclerosis exacerbation' and 'multiple sclerosis flare' fall within the same preferred term 'multiple sclerosis relapse'. For this study, the consistency in the reporting of AEs was assessed by comparing the AEs on the preferred term level. The AEs were also grouped according to the highest level of the MedDRA® hierarchy, namely the System Organ Class level.

In addition, various characteristics of the reported AEs were assessed as follows.

» *Attention*: An assessment was made of where in the text the authors described the AE. For the EPAR, it was assessed whether the regulators described the AE in the concluding section that reports how the benefits are weighted against the risks. For the scientific publications, it was assessed whether researchers described the AE in the abstract, main body of the text, a table and/or an appendix.

» *Seriousness*: Adverse events were categorized as a SAE if the authors specifically described the AE as being serious or if an AE was listed on the important medical events list of the EMA (12). An SAE is an AE that results in death, is life-threatening, requires hospitalization or prolongs

existing hospitalization, results in persistent or significant disability, or is a birth defect. This definition is also included in the guidelines for scientific publications.

» *Regulatory importance*: Adverse events were categorized as regulatory important if regulators included these as important risks in the risk management plan (RMP). A separate chapter of the EPAR describes the RMP, including the important identified risks. Regulators include safety issues as important identified risks in the RMP if these have been causally associated with the product, should be further characterized after marketing approval, and are likely to have an impact on the benefit–risk balance (13). As the EMA introduced RMPs in 2005, this information could not be included for the products authorized prior to 2005.

Data analysis

Whether the EPAR and scientific publication report consistently on AEs for the same biological was assessed by comparing these on the preferred term level. In the EPAR, when the authors referred to a pooled analysis of data, it was considered to be consistently reported if the AE was reported in at least one of the scientific publications. The frequencies of AEs that were consistently reported in both the EPAR and scientific publication, those that were only reported in the EPAR, and those that were only reported in the scientific publication were calculated.

Relative risks (including 95% confidence intervals) were calculated to assess the association of the characteristics of the AE described above and the consistency in reporting of the AE in both the EPAR and scientific publication.

Statistical analysis was performed using R statistical software version 3.6.0 (R Core Team, Vienna, Austria).

RESULTS

As of 31 December 2018, the EMA had approved nine biologicals for the treatment of MS. From these nine products, one [Extavia® (interferon-b-1b)] was excluded from the analysis as the company used the same dossier of the already available Betaferon® for the marketing approval. Although the company has taken Zinbryta® (daclizumab) off the market in March 2018, it was included in the analysis as only the information available at the time of regulatory approval was taken into account. As a result, eight biologicals were included in this study (Table 1). For all the products, the results of the phase 3 clinical trials were published in the scientific literature.

Consistency in reporting of AEs

The EPARs and/or the scientific publications reported 707 AEs. A comparable number of different AEs was reported for the interferons Avonex® (n = 23), Rebif® (n = 38) and Betaferon® (n = 33), whereas a considerably higher number was reported for the peginterferon product Plegridy® (n = 103). For the monoclonal antibodies, the number of AEs ranged from 108 for Ocrevus® to 174 for Lemtrada®.

Overall, the proportion of AEs consistently reported in both the EPAR and scientific publication was 35%. Amongst the interferons, the proportion ranged from 27% for Betaferon® to 35% for Avonex® (Fig. 1). For the monoclonal antibodies, the proportion of AEs consistently reported in both

Table 1. Biologicals authorised by the EMA for the treatment of MS included in this study.

Product name	Active substance	Year of EMA approval	Number of trials supporting the approval	Mechanism of action
Betaferon®	interferon-β-1b	1995	1	immunomodulating cytokine
Avonex®	interferon-β-1a	1997	1	immunomodulating cytokine
Rebif®	interferon-β-1a	1998	1	immunomodulating cytokine
Tysabri®	natalizumab	2006	2	anti-α4-integrin
Lemtrada®	alemtuzumab	2013	2	anti-CD52
Plegridy®	peginterferon-β-1a	2014	1	immunomodulating cytokine
Zinbryta®	daclizumab	2016 *(withdrawn 2018)*	2	anti-CD25
Ocrevus®	ocrelizumab	2018	3	anti-CD20

information sources ranged from 29% for Tysabri® to 42% for Zinbryta® (Fig. 1). Of the 707 reported AEs, 222 AEs (31%), of which 116 were SAEs, were only described in the EPAR and not in the scientific publication. Accordingly, a total of 239 AEs (34%), of which 123 were SAEs, were described only in the scientific publication. Whether more AEs were described in either the EPAR or the scientific publication differed per product. For example, for Plegridy®, 63% of the AEs were described only in the EPAR, whereas for Tysabri®, 53% of the AEs were described only in the scientific publication.

Of the 222 AEs that were described only in the EPAR, 35 (16%) were described in the section discussing the benefit–risk balance. Of the 239 AEs described only in the scientific publication, four AEs (2%) were described in the abstract of the scientific publication. The AEs were most often described in a table (50%) or the text (35%).

Serious AEs were significantly ($P < 0.05$) more often consistently reported in both the EPAR and scientific publication compared to non-serious AEs (38% vs. 30%, relative risk 1.23, 95% confidence interval 1.00–1.52) (Table 2). Also, AEs that regulators classified as important identified risk were significantly more often consistently reported in both documents compared to those that authorities did not classify as such (49% vs. 30%, relative risk 1.65, 95% confidence interval 1.34–2.03).

Nature of the reported AEs

In line with the known safety profile of the products, most AEs were infections and infestations (n = 145, 21%), followed by investigations (n = 94, 13%) and general disorders and administration site conditions (n = 70, 10%). For these categories, the consistency in reporting of the AEs ranged from 39% for infections and infestations to 49% for general disorders and administration site conditions.

The pattern of reporting SAEs in specific categories differed per product. For Avonex®, Betaferon® and Rebif®, it was not possible to observe any differences as a limited number of SAEs were reported. For Plegridy®, it was observed that five SAEs, classified as neoplasms benign, malignant and unspecified (including cysts and polyps), were only described in the EPAR. For the monoclonal antibodies, additional SAEs were reported in the EPAR and scientific publication that were related to the mechanism of action (i.e. infections and infestations) besides the SAEs that were reported in both documents. However, it was also observed that SAEs in specific categories

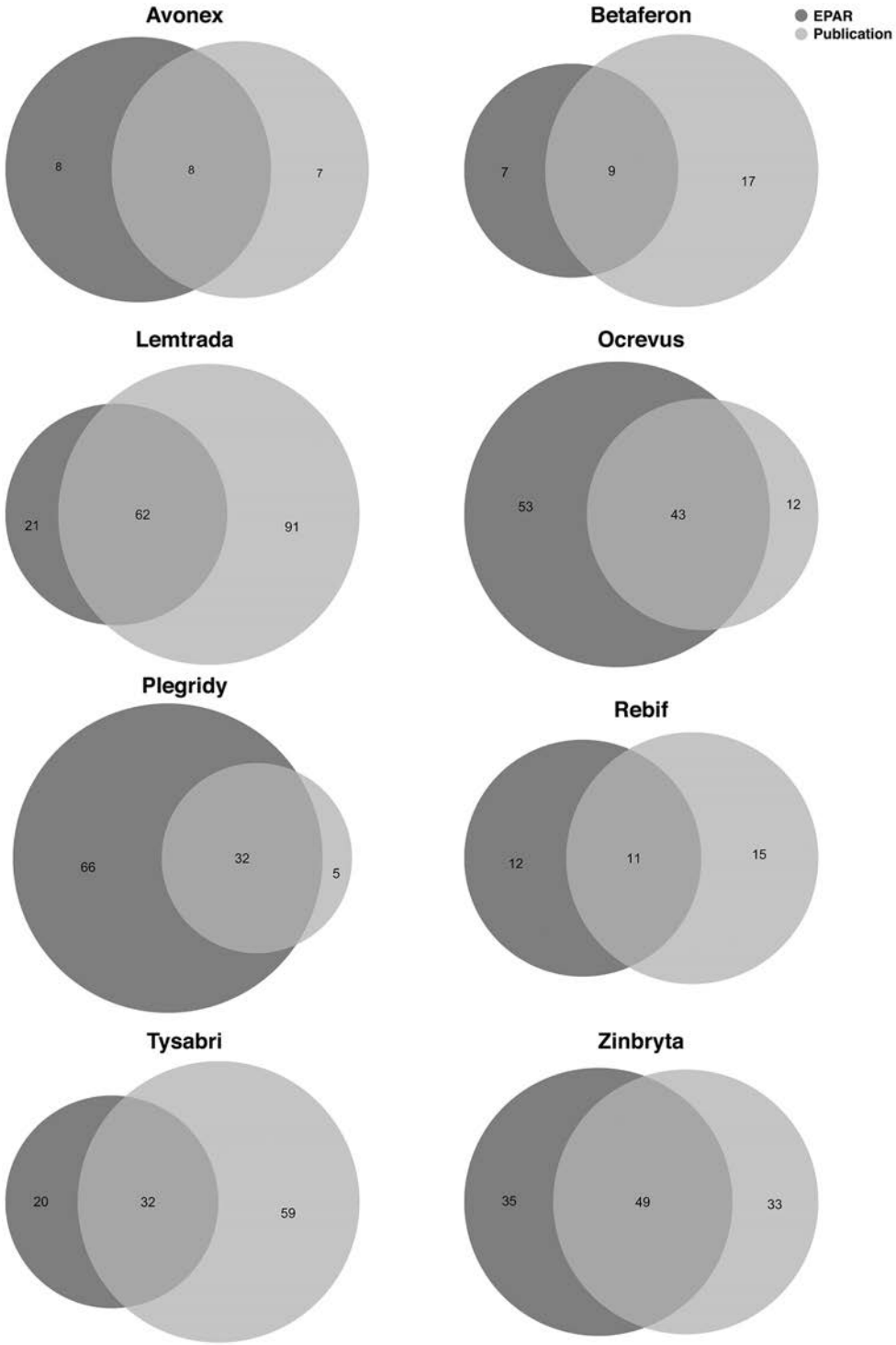


Figure 1. Venn diagrams displaying the number of AEs that were described in the EPAR and scientific publication.

Table 2. Consistency in reporting of the adverse events stratified by seriousness and regulatory importance of the adverse event.

	Consistently reported in both EPAR and scientific publication (n (%))	Relative risk (95% CI)
Serious adverse event (n=386)	147 (38%)	1.23 (1.00-1.52)
Non-serious adverse event (n=321)	99 (30%)	Reference
Adverse event classified as important identified risk (n=188)	92 (49%)	1.65 (1.34-2.03)
Adverse event not classified as important identified risk (n=425)	126 (30%)	Reference

(e.g. vascular disorders, neoplasms benign, malignant and unspecified) were only described in either one of the documents.

DISCUSSION

The current study provided a comparison of AEs reported in EPARs and scientific publications. Overall, approximately one-third of the AEs was consistently reported in both the EPAR and scientific publication, one-third in the EPAR only, and one-third in the scientific publication only. The results indicate ample discordance in the reporting of AEs between EPARs and scientific publications. However, the AEs that were reported in the EPAR or scientific publication only were, in general, not described in the most important sections of the documents, i.e. abstract or benefit–risk section. Also, SAEs and events that regulators classified as important identified risks were more often consistently reported. Therefore, both documents probably reflect the safety information that is key to the benefit–risk of the product and clinical decision making, whereas a complete overview of the AEs is lacking. This might have implications for the information presented in the clinical guidelines, including the guidelines for treatment of MS, as these are mainly based on the information that is described in the scientific publications (2). It is recommended that information from the regulators be incorporated during the development of clinical guidelines. However, the EPAR may also not reflect the complete safety profile of the product, as approximately one-third of the AEs was only reported in scientific publications. As the EPAR is a reflection of the assessment procedure, the regulators may have given specific attention to AEs that were of major concern during the assessment.

The proportion of AEs that was consistently reported was comparable amongst the products. However, whether the proportion of AEs reported in either one of the documents was higher for the EPAR or scientific publication differed per product. When looking into the nature of the AEs that were only reported in one of the documents, it was observed that these were mostly in line with the consistently reported AEs and the AEs directly linked to the mechanism of action. However, it was also observed that for some products the authors did not report on a specific type of AE, whereas the authors of the other information source did.

In line with previous studies that compared information from EPARs with scientific publications, there are differences in the information provided by the regulators and the authors of scientific publications. However, the proportion of safety information missing in the scientific publications was lower in our study compared to a previous study that performed a high-level comparison (comparing specific AEs for insomnia medication) of safety data, which reported missing safety data in eight of the 15 scientific publications (6). Also, a study that assessed reporting of SAEs in scientific publications of antidepressants found that 63% of the scientific articles did not mention any SAEs (7). These differences may be explained by the difference in the nature of the products that were included as, for example, more SAEs are associated with monoclonal antibodies used for treating MS than with the use of insomnia medication. Given these differences and as it was observed that the pattern of reporting of AEs between EPARs and scientific publications differed per product, the results may not be generalizable to other (types of) products.

For this study, all AEs that were reported at least once were considered for the included biologicals in the EPARs or scientific publications. As a causality assessment on the AEs was not performed, AEs were included that may not have been associated with the product. Also, the extraction of the AEs from the text might have been sensitive to interpretation in some cases where the authors did not specifically state whether the AE had been reported for the product under study or whether the AE was considered to be serious. However, this was minimized through consensus amongst the authors on the interpretation of different scenarios reported in the EPARs and scientific publications.

An in-depth comparison of AEs reported in the two information sources is provided and these data are put into perspective. Also, several studies considered the information from regulators as the reference information source. However, within this study it is shown that scientific publications also contribute to a complete overview of the AEs. These observations need further research on how to align the information in both sources more consistently.

Substantial discordance was observed in the AEs reported on the same phase 3 trials of biologicals for MS in information originating from regulators (described in the EPAR) and the scientific community (described in scientific publications). To support optimal clinical decision making, healthcare professionals and patients should consider both documents.

REFERENCES

1. European Medicines Agency. Zinbryta Article-20 referral - PRAC assessment report. https://www.ema.uropa.eu/en/documents/referral/zinbryta-article-20-referral-prac-assessment-report_en-0.pdf. Accessed 13/2/2019.
2. Montalban X, Gold R, Thompson AJ, Otero-Romero S, Amato MP, Chandraratna D, et al. ECTRIMS/EAN Guideline on the pharmacological treatment of people with multiple sclerosis. Mult Scler. 2018;24(2):96-120.
3. Papathanasiou P, Brassart L, Blake P, Hart A, Whitbread L, Pembrey R, et al. Transparency in drug regulation: public assessment reports in Europe and Australia. Drug Discov Today. 2016;21(11):1806-13.
4. Beijers L, Jeronimus BF, Turner EH, de Jonge P, Roest AM. Spin in RCTs of anxiety medication with a positive primary outcome: a comparison of concerns expressed by the US FDA and in the published literature. BMJ Open. 2017;7(3):e012886.
5. de Vries YA, Roest AM, Beijers L, Turner EH, de Jonge P. Bias in the reporting of harms in clinical trials of second-generation antidepressants for depression and anxiety: A meta-analysis. Eur Neuropsychopharmacol. 2016;26(11):1752-9.
6. Mattila T, Stoyanova V, Elferink A, Gispen-de Wied C, de Boer A, Wohlfarth T. Insomnia medication: do published studies reflect the complete picture of efficacy and safety? Eur Neuropsychopharmacol. 2011;21(7):500-7.
7. Serebruany VL. Discrepancies in the primary PLATO trial publication and the FDA reviews. Int J Cardiol. 2014;172(1):8-10.
8. Turner EH, Knoepflmacher D, Shapley L. Publication bias in antipsychotic trials: an analysis of efficacy comparing the published literature to the US Food and Drug Administration database. PLoS Med. 2012;9(3):e1001189.
9. Lv JW, Chen YP, Zhou GQ, Liu X, Guo Y, Mao YP, et al. Do Published Data in Trials Assessing Cancer Drugs Reflect the Real Picture of Efficacy and Safety? J Natl Compr Canc Netw. 2017;15(11):1363-71.
10. Tramacere I, Del Giovane C, Salanti G, D'Amico R, Filippini G. Immunomodulators and immunosuppressants for relapsing-remitting multiple sclerosis: a network meta-analysis. Cochrane Database Syst Rev. 2015(9):Cd011381.
11. MedDRA Maintenance and Support Services Organization. Medical Dictionary for Regulatory Activities. https://www.meddra.org/. Accessed 4/3/2019.
12. European Medicines Agency. Inclusion/exclusion criteria for the "Important Medical Events" list. https://www.ema.europa.eu/en/documents/other/eudravigilance-inclusion/exclusion-criteria-important-medical-events-list_en.pdf. Accessed 4/3/2019.
13. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP): Module V – Risk management systems (Rev 2) https://www-ema-europa-eu.proxy.library.uu.nl/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-module-v-risk-management-systems-rev-2_en.pdf. Accessed 4/3/2019.

5

GENERAL DISCUSSION

INTRODUCTION

In this thesis, regulatory and clinical insights in post-marketing safety learning for biologicals have been provided. Post-marketing safety learning is key in the drug's life-cycle (1). At the time of marketing approval, the knowledge about the safety profile of a new drug is primarily based on the results of the preclinical and clinical studies supporting approval. Uncertainties regarding the safety profile remain at the time of approval, as the safety of a drug has usually been studied in only a limited number of persons, for a limited duration, and/or in a selected population that usually differs from the clinical practice population (2). The knowledge that is available at approval is summarized in the European public assessment report (EPAR), which forms the basis for the information provided in the product information. The EPAR also includes a summary of the risk management plan (RMP), which describes the safety profile's uncertainties that must be further characterized in the post-marketing phase (3). These uncertainties can include potential safety concerns that were, for example, observed in the preclinical studies or may be related to specific contexts of use (e.g., pregnant women). These can be further characterized in post-marketing safety studies that can make use of a variety of study designs and systems (e.g., a clinical trial or an observational study using [existing] registries or large population-based databases). In addition, post-marketing safety learning is directed towards timely detection of previously unknown adverse drug reactions (ADRs). These can be detected through routine pharmacovigilance activities, such as collecting and analyzing suspected ADRs reported by patients and healthcare professionals, which should be performed for all drugs (3).

As described in *Chapter 1*, post-marketing safety learning is of specific interest for biologicals because biologicals currently represent approximately 30% of the newly launched active substances worldwide (4). Moreover, biologicals have been a challenge in terms of post-marketing safety learning over recent years given their differences with small-molecule drugs. Biologicals are produced by or extracted from a biological source, which distinguishes biologicals from small-molecule drugs as small molecules are generally produced through chemical synthesis and have relatively simple structures that can be adequately characterized. Therefore, immunogenicity reactions are more pronounced for biologicals than for small molecules due to the protein structure and formulation of biologicals. Moreover, the product characteristics of biologicals are more vulnerable to variability in the manufacturing process and formulation than that of small-molecule drugs (5). In addition to the immunogenic reactions that are more pronounced for biologicals, the nature of other ADRs is different for biologicals compared with small molecules (6-8). For biologicals, the safety issues are often related to the mechanism of action. However, the mechanism of action is not always fully elucidated, especially when the mechanism of action of the biological interferes with the immune system, which complicates the assessment of the association between the adverse event and the drug. The characterization of the safety profile of biologicals is further complicated if the symptoms of the ADRs mimic the underlying disease. This was the case for natalizumab and daclizumab, which are used to treat multiple sclerosis, a disease exerting neurological symptoms, and this complicated the detection of progressive multifocal leukoencephalopathy (PML) and encephalitis as safety events related to the use of these therapies (9, 10). Furthermore, classifying

ADRs of biologicals according to the typical distinction between type A and type B ADRs is difficult. Type A ADRs are considered to be related to the pharmacological effect of the drug and are dose dependent and common, whereas type B ADRs are classified as being unexpected and uncommon and include a variety of immunological reactions. However, for biologicals, type B ADRs are to be expected given their characteristics (11). A consequence of these challenges is that, for biologicals, the number of uncertainties regarding the safety profile at the time of approval is higher in comparison with small molecules, and in the post-marketing phase of biologicals the product information is more often updated to reflect new information regarding the safety profile (12, 13).

Given that biologicals represent an important group of drugs that are accompanied by challenges to post-marketing safety learning, various aspects of such learning for biologicals have been studied in previous research and PhD theses from our group. These studies assessed the dynamics of safety learning for both unexpected ADRs and uncertainties regarding the safety profile at approval and the characterization of specific safety issues in the post-marketing phase, including the pharmacovigilance tools used for this purpose. However, post-marketing safety learning for biologicals remains an important field to explore and is continuously evolving, with the introduction of the new European pharmacovigilance legislation in 2012 as an important regulatory milestone (14). This new legislation included, among other things, the establishment of the pharmacovigilance risk assessment committee (PRAC) of the European Medicines Agency (EMA) and the possibility of imposing safety studies in order to strengthen and rationalize post-marketing safety learning. Furthermore, a specific guideline on good pharmacovigilance practices for biologicals was introduced in 2016 (15). In addition, specific pathways that facilitate earlier access for patients to drugs were introduced that inherently increase uncertainties regarding the safety profile in the post-marketing phase, because less comprehensive data is required for these drugs at the time of approval. Finally, with the continuous introduction of drugs with new mechanisms of action, including biologicals such as PD-1/PD-L1 inhibitors and IL-17 inhibitors, additional challenges are brought to post-marketing safety learning for biologicals.

In this thesis, regulatory and clinical insights in post-marketing safety learning for biologicals have been provided by studying the characterization of specific safety issues, dynamics in post-marketing safety learning, and safety information from regulatory and clinical sources. Within this general discussion, several aspects of the findings from prior chapters will be put into a broader perspective by discussing 1) dynamics in the indication of use and dosing information for post-marketing learning; 2) challenges related to detecting and classifying ADRs; 3) optimizing post-marketing safety learning through cross-learning among drugs with commonalities in the mechanism of action; and 4) closing the gap between safety information from regulatory and clinical sources.

DYNAMICS IN THE INDICATION OF USE AND DOSING INFORMATION FOR POST-MARKETING LEARNING

Before a new drug can be approved, the benefit–risk balance in the indication under assessment should be favorable on a population level. This benefit–risk balance is, however, dependent on

the context in which the drug is used. In the clinical studies supporting approval, the drug would have been studied in a specific context, which may have been different to that in which the drug is used in clinical practice. For example, it has been demonstrated that only approximately 55% of the patients treated with empagliflozin, a sodium-glucose cotransporter 2 used for the treatment of diabetes type 2, in clinical practice would have been eligible for inclusion in the clinical trials (16). The reasons for this include concurrent use of specific glucose-lowering drugs and the presence of comorbidities (16). The clinical context in which the drug is used also changes over time. The population for which the drug is indicated can change over time when the efficacy and safety has been studied in an additional population and the use of the drug in that population is approved by regulatory authorities and/or included in the clinical treatment guidelines. Moreover, when new information becomes available that shows that the drug is not effective or cannot be safely used in a certain population (i.e., the benefit–risk has become negative), the population in which the drug is approved to be used is restricted. A high number of changes in the user population has been seen for the TNF-α inhibitors and PD-1/PD-L1 inhibitors. The TNF-α inhibitor adalimumab was, at the time of approval in 2003, indicated for the treatment of adult patients with moderate to severe, active rheumatoid arthritis when the response to disease-modifying anti-rheumatic drugs, including methotrexate, had been inadequate (17). As of December 2020, the approved indication of adalimumab has been extended 20 times in the European Union and adalimumab is currently indicated for the treatment of multiple auto-immune diseases in both adults and pediatrics (18). A similar pattern was seen for the PD-1/PD-L1 inhibitors. These were, at first, indicated for the treatment of advanced melanoma. As the PD-1/PD-L1 inhibitors have an immune-stimulating mechanism of action and melanoma is considered to be an immunogenic tumor, the first target population included melanoma patients (19). Soon after, however, the PD-1/PD-L1 inhibitors were shown to be effective for other cancer types, and they are currently approved for the treatment of a wide variety of these, including non-small-cell lung cancer, urothelial carcinoma, Hodgkin's lymphoma, and renal cell carcinoma.

In addition to the patients' characteristics that determine the drug's safety profile, the recommended dose plays an essential role in the context: according to Paracelsus, "Poison is in everything, and no thing is without poison. The dosage makes it either a poison or a remedy." Finding the optimal dose is, however, challenging. This is illustrated by the uncertainties about the optimal dose that remained at the time of ipilimumab approval (20). Ipilimumab was approved with a recommended dosing of 3 mg/kg every three weeks, although it was not fully known whether this dose induced the maximum immune activating effect. Therefore, in the post-marketing phase, a clinical trial was performed to evaluate the differences in efficacy and safety between the 3 mg/kg and 10 mg/kg dosing regimens (21). The results of this study showed that the 10 mg/kg dosing regimen resulted in a significant increase in overall survival; however, more (serious) adverse events were experienced by the patients. Eventually, it was concluded that the dosing regimen of 3 mg/kg had the optimal benefit–risk balance, and this therefore remained the recommended dosing.

Finding the optimal dose may be especially challenging for biologicals, as it is more difficult to predict their clinical effects from non-clinical data than it is for small molecules. Specifically, immune reactions such as hypersensitivity reactions and the formation of antidrug antibodies are

difficult to predict using animal models (22). A landmark example for the failure of the predictability of preclinical effects is the case of the monoclonal antibody TGN-1412, a novel anti-CD28 antibody. When first tested in humans in 2006, a cytokine storm was observed, eventually resulting in multi-organ failure (23). This cytokine storm had not been seen in the animal studies. Furthermore, finding the optimal dose for monoclonal antibodies is complicated by the complex relationship between the concentration and the response (24). In *Chapter 3.1*, we assessed whether these challenges for dose finding of biologicals are translated into changes in the dosing recommendations in the product information during the post-marketing phase. We expected that more uncertainties in the pre-approval process of biologicals could lead to more changes in the dosing information. The changes could be prompted by new information on efficacy over the drug's life-cycle, emerging safety issues related to the dose, or changes that are related to improving convenience for the patient. The results of our study, however, indicated that the dosing recommendations for the initial indication in the product information was changed for only 11% of the biologicals within a median of four years after approval, which is in line with the incidence reported for small molecules. Moreover, it was demonstrated that the dose was rarely reduced for safety reasons. Although it is reassuring that the dosing information was rarely reduced due to emerging safety issues, specific attention can be given to optimizing the dose from safety and patient convenience perspectives. From a safety perspective, optimizing the dose could be considered as a risk minimization measure. For example, multiple risk minimization measures are currently in place to reduce the risk of PML in patients treated with natalizumab. These measures include recommendations of patient monitoring through routinely performing magnetic resonance imaging scans and informing patients about the early symptoms of PML (25). These factors contribute to the risk reduction; however, reducing the dose can also contribute to this. A study found that the risk of PML could be significantly reduced by extending the dosing interval from the currently recommended interval of four weeks to approximately six weeks (26, 27). Although the regulatory authorities concluded that the dosing information would not be updated, as the efficacy of the extended dosing interval was not fully established, the information on the extended dosing interval is included in the warnings and precautions section of the product information. In addition to this risk minimization measure, patients' convenience is improved when extending the dosing interval, given that the burden for patients is reduced.

In conclusion, the challenges for dose finding for biologicals and the higher level of uncertainties at the time of approval did not result in more frequent safety related changes or dose reductions due to safety reasons in the regulatory dosing information compared with small molecules. However, it is recommended that for biologicals, reducing the dose/dosing interval is increasingly considered to be a risk minimization measure or an improvement in patient convenience. As the clinical context of use is continuously changing over time, with primarily the user population being extended, these considerations of reducing the dose/dosing interval should be taken into account for all user populations.

CHALLENGES RELATED TO DETECTING AND CLASSIFYING ADRS

In different phases of the drug's life-cycle (i.e., clinical trial and post-marketing phase), the safety profile of a drug in patients is characterized, which involves detecting and classifying ADRs. For biologicals, specific challenges are related to the detection and classification of ADRs.

The detection of ADRs for biologicals is hindered by different factors. First, the signs and symptoms experienced by the patients should be recognized as potential ADRs. When the symptoms associated with the ADRs are non-specific, such as in the case of nausea, fatigue, and asthenia, it can be difficult to distinguish between an ADR and symptoms related to the disease, especially when the drug is used for the treatment of severely ill patients. Moreover, the clinical manifestation of an ADR can differ among patients with, for example, hypersensitivity reactions that can in some patients manifest as rash and in others as fever. Furthermore, several ADRs associated with biologicals mimic the clinical expression of the disease for which the drug is used. This is, for example, seen for the biologicals used for the treatment of multiple sclerosis. For daclizumab, the early symptoms of the ADR encephalitis include aphasia, confusion, and disorientation, which are similar to the symptoms associated with a serious multiple sclerosis relapse (9). For natalizumab, the detection of PML associated with natalizumab was also challenged by the PML symptoms mimicking the disease (10). Moreover, the indication of use may challenge the detection of ADRs. For biologicals that suppress the immune system, it is acknowledged that they can contribute to an increased risk of malignancies due to their immunosuppressive action. However, the autoimmune diseases for which the biologicals are used are also associated with an increased risk of malignancies. This difficulty is illustrated by the conflicting results that are available regarding the association between treatment with TNF-α inhibitors and malignancies, for which some studies indicated an increased risk whereas others did not report an association (28-30). Several strategies have previously been described that could facilitate the detection of ADRs for biologicals. These include the early dissemination of information about potential safety signals to facilitate the recognition and reporting of suspected ADRs and keeping an open mind to the unexpected. Additionally, the knowledge and experiences of the patients can play an important role in detecting ADRs. Patients are the first link in the chain of detecting an ADR, as they are the ones who experience the signs and symptoms. In the Netherlands, since 2003, patients have been able to report their suspected ADRs to the national pharmacovigilance center and this has been shown to play an important role in the detection of, for example, the signal of aggression associated with treatment with selective serotonin reuptake inhibitors (31, 32). The role of the patients could be intensified in the detection of ADRs, because patients may be able to distinguish between the symptoms associated with the disease and ADRs, having possibly been diagnosed with the disease for several years and experienced relapses. The systematic provision of information by patients is beneficial, especially shortly after the approval of a new drug. This should be facilitated by the knowledge and open minds of healthcare professionals and be captured by structured recording of the symptoms in order to enable the detection of any differences in them. Although this information is increasingly being collected as part of clinical trials and in clinical practice to inform clinical decision making, also for the detection of ADRs in the post-marketing phase this information could be of added value

(33, 34). Moreover, for known ADRs the detection can be optimized in clinical practice. Several ADRs are detected through laboratory monitoring, including the thyroid disorders associated with PD-1/PD-L1 inhibitors described in *Chapters 2.2 and 4.1*. The detection of these ADRs can be facilitated by automatically linking the requests for the laboratory tests to the drug records at the time of prescribing and therefore ensuring that the laboratory tests are performed. A prerequisite for this is that the timing of the measurements is specified in the product information. Currently, the product information states only that thyroid levels should be monitored periodically, without specifying the exact interval. We illustrated in *Chapter 2.2* that the thyroid disorders mainly occur in the first months of treatment, thus the current recommendations should be replaced by more specific recommendations, recommending that thyroid levels should be measured at every treatment cycle and especially during the first treatment cycles.

In addition to the challenges faced in the detection of the ADRs, the classification of the ADRs is complex. Classification of the ADRs is an important part of post-marketing safety learning, because it facilitates the comparability of safety information and the exchange of information between different stakeholders, such as between healthcare professionals, pharmaceutical companies, and regulatory authorities. Different dictionaries are place for the classification of the ADRs. Within oncology, the National Cancer Institute's common terminology criteria for adverse events (CTCAE) are applied in classifying adverse events in clinical trials (35). The CTCAE dictionary is also used to describe the severity of the ADRs in the product information, with consequent dose recommendations. Although the CTCAE dictionary contributes to the standardization of classifying ADRs, there is room for improvement, as illustrated in *Chapter 4.1*, in which we assessed how patients are classified as having thyroid disorders in the clinical trials and observational studies. We demonstrated that, in all of the 38 clinical trials, the CTCAE dictionary was used to classify the thyroid disorders. However, the classification of thyroid disorders described in the CTCAE dictionary (i.e., hyperthyroidism and hypothyroidism) was non-specific, as the levels at which the thyroid hormones meet the classification of hyperthyroidism and hypothyroidism were not described. Observational studies, in contrast to the clinical trials, generally included a description of the reference values of the thyroid hormones; however, the exact reference values differed between observational studies. To improve the comparability among the studies, the CTCAE dictionary and clinical practice guidelines could provide more conclusive classifications by including the reference values for the thyroid hormones. Among the clinical trials, differences were observed in the classification of ADRs as immune-related ADRs. As immune-related ADRs are the distinct ADRs associated with PD-1/PD-L1 inhibitors, standardization of this classification could improve the ability to compare ADRs among PD-1/PD-L1 inhibitors and between indications for the same PD-1/PD-L1 inhibitor. The challenges seen for the classification of (immune-related) thyroid disorders associated with the PD-1/PD-L1 inhibitors are expected not to be limited to these ADRs and drugs. The thyroid disorders can, as is the case for other ADRs such as fever and renal disorders, be objectively classified using laboratory values, and therefore a standardized classification could have been expected. Moreover, there was a window of opportunity to standardize the classification of the ADRs of the PD-1/PD-L1 inhibitors, as these were authorized shortly after each other. Therefore, the standardization of the classification of ADRs should be facilitated for all ADRs, especially for those that can be objectively classified according

to biomarkers. Regulatory authorities should play a role in the standardization of ADRs since they provide, at an early stage in drug development, advice on the conduct of the clinical trials, including advice on the protocol. This regulatory advice could include an assessment of the classifications of ADRs, thereby assuring a more harmonized classification of the ADRs. In addition to the CTCAE dictionary that is used for the description of the severity of the ADRs in the product information, the medical dictionary for regulatory activities (MedDRA) is used to classify the ADRs described in the product information. MedDRA is hierarchically structured, and, in the section that lists the ADRs, these are generally classified on the preferred-term level. On this level, for example, a distinction is made between the preferred terms "abdominal pain upper" and "abdominal pain lower" (36). The clinical significance of including multiple preferred terms in the product information may, however, be limited. This was also the underlying rationale for assessing the overlap in ADRs in the product information of TNF-α inhibitors, as described in *Chapter 3.2*, on the high-level-term level. When the clinical meaning and management of the ADRs is independent of the preferred-term or high-level-term level, describing ADRs as high-level terms could improve the comparability of ADRs among drugs with commonalities in the mechanism of action. Although we showed in *Chapter 3.2* that only 55% of all ADRs described for TNF-α inhibitors were described in the product information of at least two TNF-α inhibitors on the high-level-term level, this proportion would have been substantially lower had we studied the overlap on the preferred-term level, indicating that the comparability of the ADRs described in the product information is even more hampered.

In conclusion, the detection and classification of ADRs for biologicals need to be improved. First, the detection of ADRs can be facilitated by increasingly making use of the observations of the patients, especially in a context where ADRs mimic the signs and symptoms associated with the disease for which the drug is used. Furthermore, when ADRs are detected through routine laboratory tests, automatically linking the drug prescriptions to the requests for laboratory tests will facilitate the detection of the ADRs. For this, the product information should be updated to include more specific recommendations for the monitoring interval. For the classification of the ADRs, the dictionaries used to classify ADRs and clinical guidelines should be improved by using conclusive classifications, and this also facilitates the classification of ADRs in clinical practice studies. This can be given direction by regulatory authorities, which provide advice in an early stage of clinical development (i.e., protocol phase). In particular when drugs with a comparable safety profile are authorized shortly after each other, there is a window of opportunity to standardize the classification of the ADRs. Finally, the terms used to describe the ADRs in the product information are currently very specific. The clinical applicability and comparability among drugs can be improved by using slightly broader terms to describe the ADRs without impacting the clinical meaning of them.

OPTIMIZING POST-MARKETING SAFETY LEARNING THROUGH CROSS-LEARNING AMONG DRUGS WITH COMMONALITIES IN THE MECHANISM OF ACTION

When a patient uses a drug and experiences any unexpected symptoms, these are considered to be adverse events. However, these adverse events are not necessarily related to the drug used.

It should be assessed whether the adverse event is related to the drug and would therefore be considered to be an ADR. Different criteria play a role in this assessment, such as a plausible time to onset, whether the symptoms disappear or reduce when the treatment is stopped, and whether no other factors can explain the occurrence of the event. Moreover, an important criterion is whether there is a mechanistic plausibility for the occurrence of the ADR. In numerous cases, the underlying mechanism of occurrence of the ADR has been elucidated and linked to the mechanism of action of the drug. For example, treatment with the VEGF inhibitor bevacizumab may cause impaired wound healing because VEGF plays an important role in the wound-healing cascade (37, 38). For tuberculosis infections associated with TNF-α inhibitors, the involvement of TNF-α was clearly determined, and tuberculosis infections are therefore thought to occur independently of the specific TNF-α inhibitor used (39). In addition, the cardiotoxic effects that are commonly seen in patients treated with HER-2 inhibitors (e.g., pertuzumab and trastuzumab) are explained by the involvement of HER-2 in cardiac cells (40-42). In these examples, the association between the mechanism of action and the occurrence of the ADR was fully elucidated. However, as previously described in *Chapter 1*, the mechanism of action of biologicals, including the potential interference with the immune system, is not always fully elucidated, and therefore the determination of the association with the ADR is confounded. With extensive research being performed, the pathophysiology of multiple auto-immune diseases is being increasingly unraveled. Moreover, emerging safety issues can provide additional insights into the pathways that are involved. For example, the auto-immune reactions seen in patients treated with daclizumab have provided insight into the potential role of the innate immune system in the treatment of auto-immune diseases (43). Nevertheless, full knowledge of the mechanism of action of biologicals and the potential pathways affected remains a challenge that post-marketing safety learning remains to face in the (near) future.

Despite the limited knowledge about the exact mechanism of action for a number of biologicals, there may be opportunities to optimize mechanism-of-action-related post-marketing safety learning. For this, safety information from one biological could be transposed to another biological with commonalities in the mechanism of action. This cross-learning among drugs with commonalities in their mechanisms of action may be especially suited to biologicals. First, ADRs of biologicals are, in addition to immunogenic reactions, mainly related to the mechanism of action. Secondly, biologicals are authorized in contexts for which limited information is available, which is, for example, the case for orphan drugs or drugs authorized through specific regulatory pathways aimed to facilitate early drug access for patients. Thirdly, when the outcome is considered to be rare, cross-learning among drugs with commonalities in the mechanism of action could be of value. Therefore, two strategies are proposed to facilitate mechanism-of-action-related post-marketing safety learning through cross-learning among drugs with commonalities in the mechanism of action: 1) cross-learning can be facilitated by classifying drugs according to commonalities in the mechanism of action, which can be performed to provide input for regulatory safety information as well as for observational studies performed in clinical practice; and 2) current signal detection methods can be complemented by new methods, with incorporation of data from multiple information sources.

Cross-learning among drugs with commonalities in the mechanism of action

To facilitate post-marketing safety learning among drugs with commonalities in the mechanism of action, drugs can be mechanistically classified dependent on the intended purpose. Classifying drugs on a target level (e.g., TNF-α) could directly provide input on the safety learning among drugs with the same target. Classifying drugs on a higher level (e.g., influence on the immune system) could generate hypotheses for further safety learning.

On a target level, for example TNF-α, it is expected that drugs cause similar ADRs. Safety information from one TNF-α inhibitor could therefore provide input for the safety learning of another TNF-α inhibitor. We assessed whether this was the case in *Chapter 3.2*, by studying the overlap in ADRs described in the product information of the TNF-α inhibitors. In *Chapter 3.2*, we illustrated that the overlap in ADRs described in the product information of the TNF-α inhibitors is limited: 45% of all ADRs described for the TNF-α inhibitors was described in the product information of only one TNF-α inhibitor. Moreover, acquired knowledge about the ADRs associated with TNF-α inhibitors is not fully transferred when additional TNF-α inhibitors are approved: only 39% of the ADRs that were identified prior to the approval of non-first-in-class TNF-α inhibitors were described in the product information at the approval of the non-first-in-class TNF-α inhibitor. The study described in *Chapter 3.2* builds on existing knowledge showing that regulatory cross-learning among drugs with the same mechanism of action is limited. Stefansdottir et al. illustrated that 40% of the ADRs were described in the product information of both the first- and second-in-class drugs (based on indication, mechanism of action, and structure of the drug) (44). The results of these studies indicate that there is a potential in optimizing post-marketing safety learning based on information gathered through classifying drugs according to specific characteristics.

Optimizing post-marketing safety learning among drugs with commonalities in the mechanism of action could improve the provision of safety information for healthcare professionals and patients. Although we did not study the implications for clinical practice of the limited overlap in ADRs described in the product information, as illustrated in *Chapter 3.2*, it could be expected this may hamper the adequate provision of safety information. For example, at the time of approval of durvalumab, a PD-L1 inhibitor, experience with the safety profile of comparable drugs was gained because other PD-1/PD-L1 inhibitors had been used in clinical practice for several years. The regulatory authorities therefore concluded that the extensive measures to minimize the risk of immune-related ADRs were not considered necessary for durvalumab (45). At the time of approval, however, several ADRs, including the serious neurological disorders myasthenia gravis and Guillain-Barré syndrome, were not described in the product information of durvalumab, whereas these were known to be associated with other PD-1/PD-L1 inhibitors. The implications for healthcare professionals were likely to be limited given that they were familiar with the safety profile of the PD-1/PD-L1 inhibitors. However, for patients that are not familiar with these drugs there may have been implications, as they were not informed through the patient leaflet about these ADRs and the early symptoms associated with it. Eventually, approximately one and a half to two years after approval, the product information was updated to include these ADRs (46). Patients were therefore not adequately informed about the ADRs for a substantial amount of time. This lack of

timely information about the safety of drugs could be improved in the future through cross-learning among drugs with commonalities in the mechanism of action.

In order to optimize the knowledge transfer among drugs with commonalities in the mechanism of action, potential class effects should be assessed at different points in the drug life-cycle. First, at the time of the approval of a new drug, the safety profile of drugs with commonalities in the mechanism of action (for example, with the same target) may have developed as these have already been used in clinical practice. As shown in *Chapter 3.2*, at the time of the approval of the last two TNF-α inhibitors (certolizumab and golimumab), a total of 238 out of 318 ADRs were already described in the product information of the other TNF-α inhibitors that had been used in clinical practice for multiple years. Currently, in the RMP, there is an opportunity to reflect on the safety issues observed for other drugs with the same mechanism of action. When the safety issues observed for in-class products are considered to be key to the benefit–risk balance and should be further characterized in the post-marketing phase, these are included in the RMP as important potential risks (3). For example, at the time of the approval of isatuximab, a CD38 antibody indicated for the treatment of multiple myeloma, viral reactivation was included in the RMP as an important potential risk because daratumumab, another CD38 antibody indicated for the treatment of multiple myeloma, has a known risk of hepatitis B reactivation (47). However, for the optimal provision of safety information, the ADRs associated with these comparable drugs should not only be listed in the RMP but also be included in the product information at the time of approval. Moreover, in the post-marketing phase, cross-learning among drugs with commonalities in the mechanism of action should increasingly occur. In the post-marketing phase, there are procedures (i.e., signal and referral procedures) in place to evaluate specific safety issues for a class of drugs and update the product information for all the drugs at once. However, the lag time between the first description of an ADR in the product information to uptake of this ADR in the product information of another drug can be long. As illustrated in *Chapter 3.2*, the duration of time for this cross-learning in TNF-α inhibitors was approximately three years. This is because, within the regulatory system, other procedures are also in place to update the product information to describe new ADRs. These procedures evaluate the drugs separately or on an active-substance level in the case of a single assessment of periodic safety update reports (PSURs). Although in the PSUR, in line with the RMP, safety issues that are associated with other in-class products should be reported, these are limited to specific safety issues. To facilitate the cross-learning among drugs with the same mechanism of action in the post-marketing phase, a specific section in the PSUR could be included to reflect on the ADRs that have been included in the product information of in-class products. For this, comparable text-mining tools such as that used in the study described in *Chapter 3.2* could facilitate the identification of the ADRs described in the product information of in-class drugs. The data supporting the addition of ADRs in the product information is overseen by the EMA. Therefore, the EMA is advised to provide guidance on the ADRs that should be included for the assessment in the post-marketing phase as well as at the time of approval.

There are various opportunities on how to optimize cross-learning among drugs with commonalities in the mechanism of action during all phases of drug development. In the preclinical phase, efforts are currently being made to facilitate the assessment of comparability among

biologicals with the same active substance and with the same mechanism of action. The World Health Organization has recently developed an international standard for infliximab, with which the bioactivity of different infliximab products can be compared (48). Furthermore, the European Pharmacopoeia Commission has a pilot project running to develop standardized bioassays for the TNF-α inhibitors (49). When universal methods are applied for the determination of the potency of TNF-α inhibitors, it can be assessed whether the potency is comparable among the TNF-α inhibitors and therefore whether the safety profile is expected to be mostly comparable or if there may be variation potentially leading to differences in the safety profile. This information can be used as input for the safety information provided in the product information. Moreover, besides the direct input on the safety learning among drugs with the same target in the drug's life-cycle, post-marketing safety monitoring provides opportunities for generating hypotheses based on the mechanistical commonalities. This approach was used in *Chapter 2.1*, in which we studied the potential association between the reporting of depression and suicidal ideation and the use of monoclonal antibodies, thereby classifying the monoclonal antibodies according to their influence on the immune system. For the monoclonal antibodies, a limited number of depression and suicidal ideation and behavior reports were available in VigiBase, the global database of reports of suspected ADRs. A total of 9455 reports were available for depression and 1770 for suicidal ideation and behavior. By grouping the monoclonal antibodies according to their influence on the immune system, we were able to demonstrate that the association was strongest for the monoclonal antibodies suppressing the immune system that were used for treating autoimmune diseases compared with the monoclonal antibodies not directly targeting the immune system. These results provide input for future research in which this potential risk of depression and suicidal ideation and behavior is further characterized and quantified.

Post-marketing safety learning among drugs with commonalities in the mechanism of action should be optimized to improve the adequate provision of safety information to patients and healthcare professionals. From a regulatory perspective, potential class effects (e.g., on a target level) should be considered at different points in the drug life-cycle, which could be facilitated by the EMA. These class effects should then also be included in the product information. Moreover, drugs can be classified according to other commonalities in the mechanism of action in studies performed in clinical practice to generate hypotheses for further safety learning.

Complementing existing methods for post-marketing safety learning

The exchange of safety information among drugs with commonalities in the mechanism of action could also increasingly be facilitated by applying new methods using information from established sources. One of the main established information sources used for post-marketing safety learning comprises spontaneous reporting databases. The European spontaneous reporting database EudraVigilance contained over 16.7 million reports of suspected ADRs in 2019 and contributed to 55% of the signals assessed by the PRAC (50, 51). Spontaneous reports were shown to be an important information source for the support of safety-related regulatory actions for biologicals and are predicted to remain so in the future (52, 53). Another important information source that

contributes to signal detection is the scientific literature: 13% of the signals assessed by the PRAC came from scientific literature. For example, in 2018, the PRAC recommended that dolutegravir should not be used in women seeking to become pregnant (54). This recommendation was based on the preliminary results of a study evaluating birth outcomes in babies born to women treated with dolutegravir that showed that dolutegravir appears to increase the risk of neural tube defects (55). In addition, other information sources, such as information from preclinical and clinical studies, provide input for signal detection (51). Currently, the data derived from each of the information sources is separately used as an input for the analysis of the potential safety signal. However, there is potential to integrate information from different sources to facilitate post-marketing safety learning using prediction models. For example, within the United States Food and Drug Administration (FDA), a model was developed to predict post-marketing ADRs for a new drug at the time of approval. For the model, data was used from the FDA adverse event reporting database FAERS, scientific literature, and the product information of drugs that have a similar target activity (56). The applicability of this model is not limited to the context of the United States and could also have used data from the European regulatory system, as the infrastructure is in place with, for example, the European reporting database EudraVigilance. In the prediction models, knowledge about the structure of the drugs, their mechanism of action, and the pathways involved can also be incorporated. Liu et al. developed a model to predict ADRs, combining information about the chemical properties (chemical substructures), biological properties (drug protein target, transporters, enzymes, and derived pathway information from the protein targets), indications, and other known ADRs (57). These models provide an integrated prediction of the potential association between the drug and the occurrence of an ADR, instead of separately weighing the data. Another factor that could be evaluated as part of the prediction model is the seriousness of the ADR. The seriousness of the ADR plays a role in the tradeoff between the acceptance of the association between the drug and the ADR and the uncertainties surrounding the association, as illustrated in *Chapter 3.2*. In this chapter, we showed that serious ADRs and ADRs that were classified by the regulators as important risks were included in the product information of at least two TNF-α inhibitors approximately four times more often than ADRs not classified as such. The seriousness of the ADRs can therefore be used as an additional factor when deciding on the inclusion of the ADRs in the product information based on the results of the prediction models.

By applying these strategies of combining data from different sources and making use of prediction modelling, post-marketing safety learning can increasingly be tailored towards specific safety concerns. For the further characterization of potential safety issues, alternative approaches have been proposed. Tatonetti proposed the integration of observational data from humans with laboratory experiments in model systems (58). This includes a three-step approach that starts with the detection of a potential signal through mining large observational databases. These hypotheses are then evaluated with data from another dataset and assessed for plausibility, after which the remaining hypotheses are validated through prospective experiments (58). This method was used to discover drug–drug interactions that cause QT prolongation (59). First, FAERS was searched to identify potential drug–drug interactions causing QT prolongation. The potential safety signals were then evaluated using a database containing electrocardiograms of patients treated

at a university medical center. Finally, the identified potential drug–drug interaction between ceftriaxone and lansoprazole was validated in a laboratory experiment. Thus, potential safety issues can be validated, and, by combining information from multiple sources, the causality assessment can be facilitated.

The approaches described above, together with existing methods, are potentially valuable in facilitating cross-learning among drugs with commonalities in the mechanism of action. Therefore, the regulatory applicability of these methods should be determined.

CLOSING THE GAP BETWEEN SAFETY INFORMATION FROM REGULATORY AND CLINICAL SOURCES

Physicians and other healthcare professionals (e.g., pharmacists) can consult multiple information sources to facilitate clinical decision making when treating patients. These information sources can be classified in text sources, including articles in scientific journals, clinical guidelines, and regulatory documents, electronic sources such as UpToDate and Medline, and interpersonal sources, including colleagues or consultants (60). The type of information source that is consulted to support clinical decision making depends on multiple factors. Relevance, credibility, reliability, accessibility, and usability have shown to influence the choice of the information source (60). One of the most important reasons why physicians consult information sources is to seek information on the choice of appropriate treatment for the patient. Physicians' needs when prescribing drugs include information on the indication, dose, duration of use, and safety profile of the drugs (61). As described in *Chapter 1*, post-marketing safety learning is facilitated by the adequate provision of safety information. For example, risks could be minimized when the recommended risk minimization measures are adequately taken, and, when physicians are familiar with the complete safety profile, the reporting of suspected ADRs could be facilitated. However, given the differences in the intended purpose of the sources, the information provided can vary among them.

The EPARs and scientific publications, two of the information sources that provide safety information about drugs, were assessed in the study described in *Chapter 4.2*. For this study, we compared information about the adverse events derived from the same phase three clinical trials from these two information sources. We included biologicals used for the treatment of multiple sclerosis and showed that approximately one-third of the adverse events were reported in both the EPAR and scientific publication, one-third was reported only in the EPAR, and one-third only in the scientific publication. Our study showed that neither of the two information sources provide a complete overview of the adverse events, which illustrates that there is a gap between information from regulatory and clinical sources.

The gap between information from regulatory and clinical sources is also seen when the results of the study described in *Chapter 3.1* are put into a broader perspective. We assessed only the regulatory information source for that study (i.e., product information) to investigate the number and nature of dosing changes for biologicals and illustrated that the dosing information of biologicals is rarely reduced for safety reasons. This does not always align with clinical practice where initiatives are taken to optimize the dose that may not be described in the regulatory information sources. One

factor that plays a role in optimizing the dose for biologicals in clinical practice are the costs. In 2017, biologicals comprised 37% of the net drug spending in the United States, whereas they represented 2% of all prescriptions (62). In order to reduce costs, initiatives were taken in clinical practice in 2017 to evaluate the potential of fixed dosing for monoclonal antibodies in oncology instead of the (at that time, common) weight-based dosing. A study performed by investigators of the Dutch Cancer Institute combined publicly available data from clinical trials and pharmacokinetic modeling to assess the applicability of fixed dosing of monoclonal antibodies in oncology (63). They showed that the effects of body weight on the volume of distribution and clearance were either limited or, when the impact was strong, the range of the therapeutic window was considered to be wide enough to justify fixed dosing. Therefore, it was concluded that fixed dosing for monoclonal antibodies in oncology is justified, although in some cases fixed doses are proposed for different weight ranges. Companies responsible for an important group of monoclonal antibodies in oncology, the PD-1/PD-L1 inhibitors, also illustrated that there were no clinically significant differences between weight-based and fixed dosing. Therefore, fixed dosing is currently the recommended dose described in the product information of the majority of the PD-1/PD-L1 inhibitors (64, 65). The dosing recommendations described in the product information are, however, not fully in line with the dosing applied by the Dutch Cancer Institute. For example, the recommended dose described in the product information of pembrolizumab is 200 mg every three weeks, whereas the Dutch Cancer Institute recommends a dose of 150 mg every three weeks (for patients 40–140 kg) (63, 64). The differences between the dosing information that is described in regulatory information sources and that applied in clinical practice are not limited to the fixed dosing example of PD-1/PD-L1 inhibitors. A study in a Dutch hospital compared standard and extended infusion intervals in patients treated with eculizumab for atypical hemolytic uremic syndrome (66). Given that eculizumab is costly and the burden for patients is high because eculizumab is administered every two weeks, an extended dosing interval could optimize the therapy from a patient and healthcare perspective. The study showed that extending the dosing interval to every three to four weeks may be optimal in a substantial subset of patients (66). The outcomes of this study formed the basis of the dose recommendations described in the Dutch treatment guideline for patients with atypical hemolytic uremic syndrome. Similar has been shown for the TNF-α inhibitors, where a review of data from (randomized) controlled trials revealed that, in rheumatoid arthritis patients with a low disease activity, the dose can effectively be down titrated to a dose below the dosing recommendation described in the product information (67).

The studies described in *Chapter 4.2 and 3.1* illustrate the gap between safety information in regulatory and clinical information sources, which diverges during the of the drug's life-cycle. The gap can be closed by improving the clinical applicability of regulatory information sources and incorporating information from clinical sources in the regulatory processes.

First, the gap between regulatory and clinical information sources can be closed by improving the applicability of the information provided in the regulatory information sources for clinical practice. Although for individual physicians the applicability is limited, regulatory information sources provide additional information compared with scientific publications, as illustrated by the study described in *Chapter 4.2*. Moreover, it was previously shown that EPARs provide additional

information on appropriate prescribing for older people compared with the product information (68). Furthermore, the regulatory authorities have extensively assessed all available data and may have concluded that the drug does not have a favorable benefit–risk profile for a certain population, which could be of value for the recommendations provided in the clinical guidelines. Therefore, on a non-individual level, when developing clinical guidelines, it is suggested that information from regulatory sources is incorporated. For the development of clinical guidelines, among others, systematic reviews are used that are mainly based on scientific publications (69, 70). Cochrane is exploring the possibilities of including regulatory information in their reviews (71). However, including regulatory documents in the Cochrane systematic reviews is currently only optional. One complicating factor in including regulatory information, besides the publicly availability of the information, could be the format that is used. In contrast to the reporting of harm in scientific publications, the reporting of adverse events in the EPAR is not standardized (72). To improve the clinical applicability of regulatory information sources, efforts should be made to standardize the reported information.

In addition, information from clinical sources should be incorporated more within regulatory processes to close the gap between regulatory and clinical sources. For example, observational studies such as that described in *Chapter 2.2*, in which we assessed the thyroid disorders associated with PD-1/PD-L1 inhibitor treatment, can provide additional insights into the use of the drug in clinical practice including, for example, the incidences of ADRs. Multiple barriers for the inclusion of information from clinical sources in regulatory information sources could be identified, including practical and legal barriers. However, within current pharmacovigilance tools, such as the PSUR, relevant information from clinical practice, such as published literature, is discussed. In practice, the attention given to this information can be increased.

CONCLUSION

In this thesis, regulatory and clinical insights in post-marketing safety learning for biologicals have been provided and evaluated in a broader perspective. In conclusion, the dynamics in dosing information play an important role in the context of post-marketing learning and could be optimized for risk minimization purposes and patient's convenience. The detection of ADRs for biologicals could be facilitated by intensifying the role of the patients and implementing monitor recommendations to be applied in clinical practice. For the classification of ADRs, efforts should be made to standardize them, especially those that can be objectively detected by using biomarkers. Post-marketing safety learning can be facilitated by cross-learning among drugs with commonalities in the mechanism of action by grouping them according to these commonalities and by complementing existing methods using prediction models that combine data from multiple information sources. Finally, the gap between safety information from regulatory and clinical sources can be closed by improving the clinical applicability of regulatory information sources and increasingly incorporating information from clinical practice into the regulatory system. Further studies could facilitate the optimization of post-marketing safety learning for biologicals with the recommendations made in this thesis by, for example, studying the EU applicability of implementing and validating prediction modelling for

cross-learning among drugs with commonalities in the mechanism of action, developing reference standards for the classification of ADRs detected through biomarkers, and studying preferences of clinical practice for the safety information provided in regulatory sources.

REFERENCES

1. World Health Organization. The importance of pharmacovigilance. https://apps.who.int/iris/handle/10665/42493. Accessed 1/11/2020.
2. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet. 2005;365(9453):82-93.
3. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP): Module V – Risk management systems (Rev 2) https://www-ema-europa-eu.proxy.library.uu.nl/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-module-v-risk-management-systems-rev-2_en.pdf. Accessed 15/11/2020.
4. IQVIA Institute for Human Data Science. The Global Use of Medicine in 2019 and Outlook to 2023. https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/the-global-use-of-medicine-in-2019-and-outlook-to-2023.pdf?_=1606729931964. Accessed 12/10/2020.
5. Kuriakose A, Chirmule N, Nair P. Immunogenicity of Biotherapeutics: Causes and Association with Posttranslational Modifications. J Immunol Res. 2016;2016:1298473.
6. Giezen TJ, Mantel-Teeuwisse AK, Meyboom RH, Straus SM, Leufkens HG, Egberts TC. Mapping the safety profile of biologicals: a disproportionality analysis using the WHO adverse drug reaction database, VigiBase. Drug Saf. 2010;33(10):865-78.
7. Ebbers HC, Al-Temimi E, Moors EH, Mantel-Teeuwisse AK, Schellekens H, Leufkens HG. Differences between post-authorization adverse drug reactions of biopharmaceuticals and small molecules. BioDrugs. 2013;27(2):167-74.
8. Giezen TJ, Mantel-Teeuwisse AK, Straus SM, Schellekens H, Leufkens HG, Egberts AC. Safety-related regulatory actions for biologicals approved in the United States and the European Union. Jama. 2008;300(16):1887-96.
9. European Medicines Agency. Zinbryta Article-20 referral - PRAC assessment report. https://www.ema.europa.eu/en/documents/referral/zinbryta-article-20-referral-prac-assessment-report_en-0.pdf. Accessed 25/7/2020.
10. European Medicines Agency. Tysabri: EPAR - Scientific discussion. https://www.ema.europa.eu/en/documents/scientific-discussion/tysabri-epar-scientific-discussion_en.pdf. Accessed 17/01/2021.
11. Meyboom RH, Lindquist M, Egberts AC. An ABC of drug-related problems. Drug Saf. 2000;22(6):415-23.
12. Vermeer NS, Duijnhoven RG, Straus SM, Mantel-Teeuwisse AK, Arlett PR, Egberts AC, et al. Risk management plans as a tool for proactive pharmacovigilance: a cohort study of newly approved drugs in Europe. Clin Pharmacol Ther. 2014;96(6):723-31.
13. Downing NS, Shah ND, Aminawung JA, Pease AM, Zeitoun JD, Krumholz HM, et al. Postmarket Safety Events Among Novel Therapeutics Approved by the US Food and Drug Administration Between 2001 and 2010. Jama. 2017;317(18):1854-63.
14. European Commission. Directive 2010/84/EU of The European Parliament and of the Council of 15 December 2010, amending, as regards to pharmacovigilance, Directive 2001/83/EC on the Community Code relating to medicinal products for human use. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:348:0074:0099:EN:PDF. Accessed 15/01/2021.
15. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP). Product- or Population-Specific Considerations II: Biological medicinal products. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-product-population-specific-considerations-ii_en-0.pdf. Accessed 16/12/2020.
16. Munk NE, Knudsen JS, Pottegård A, Witte DR, Thomsen RW. Differences Between Randomized Clinical Trial Participants and Real-World Empagliflozin Users and the Changes in Their Glycated Hemoglobin Levels. JAMA Netw Open. 2020;3(2):e1920949.

17. European Medicines Agency. Humira: EPAR - Scientific discussion. https://www.ema.europa.eu/en/documents/scientific-discussion/humira-epar-scientific-discussion_en.pdf. Accessed 17/01/2021.

18. European Medicines Agency. Humira: EPAR - Procedural steps taken and scientific information after authorisation. https://www.ema.europa.eu/en/documents/procedural-steps-after/humira-epar-procedural-steps-taken-scientific-information-after-authorisation_en.pdf. Accessed 17/1/2021.

19. Sanlorenzo M, Vujic I, Posch C, Dajee A, Yen A, Kim S, et al. Melanoma immunotherapy. Cancer Biol Ther. 2014;15(6):665-74.

20. European Medicines Agency. Yervoy: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/yervoy-epar-public-assessment-report_en.pdf. Accessed 16/11/2020.

21. Ascierto PA, Del Vecchio M, Robert C, Mackiewicz A, Chiarion-Sileni V, Arance A, et al. Ipilimumab 10 mg/kg versus ipilimumab 3 mg/kg in patients with unresectable or metastatic melanoma: a randomised, double-blind, multicentre, phase 3 trial. Lancet Oncol. 2017;18(5):611-22.

22. van Meer PJ, Kooijman M, Brinks V, Gispen-de Wied CC, Silva-Lima B, Moors EH, et al. Immunogenicity of mAbs in non-human primates during nonclinical safety assessment. MAbs. 2013;5(5):810-6.

23. Suntharalingam G, Perry MR, Ward S, Brett SJ, Castello-Cortes A, Brunner MD, et al. Cytokine storm in a phase 1 trial of the anti-CD28 monoclonal antibody TGN1412. N Engl J Med. 2006;355(10):1018-28.

24. Ternant D, Paintaud G. Pharmacokinetics and concentration-effect relationships of therapeutic monoclonal antibodies and fusion proteins. Expert Opin Biol Ther. 2005;5 Suppl 1:S37-47.

25. European Medicines Agency. Tysabri: EPAR - Product Information. https://www.ema.europa.eu/en/documents/product-information/tysabri-epar-product-information_en.pdf. Accessed 20/1/2021.

26. Ryerson LZ, Foley J, Chang I, Kister I, Cutter G, Metzger R, et al. Natalizumab extended interval dosing (EID) is associated with a reduced risk of progressive multifocal leukoencephalopathy (PML) than every-4-week (Q4W) dosing: Updated analysis of the TOUCH® Prescribing Program database (1988). Neurology. 2020;94(15 Supplement):1988.

27. Ryerson LZ, Foley J, Chang I, Kister I, Cutter G, Metzger RR, et al. Risk of natalizumab-associated PML in patients with MS is reduced with extended interval dosing. Neurology. 2019;93(15):e1452-e62.

28. Bongartz T, Sutton AJ, Sweeting MJ, Buchan I, Matteson EL, Montori V. Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: systematic review and meta-analysis of rare harmful effects in randomized controlled trials. Jama. 2006;295(19):2275-85.

29. Beukelman T, Haynes K, Curtis JR, Xie F, Chen L, Bemrich-Stolz CJ, et al. Rates of malignancy associated with juvenile idiopathic arthritis and its treatment. Arthritis Rheum. 2012;64(4):1263-71.

30. Dommasch ED, Abuabara K, Shin DB, Nguyen J, Troxel AB, Gelfand JM. The risk of infection and malignancy with tumor necrosis factor antagonists in adults with psoriatic disease: a systematic review and meta-analysis of randomized controlled trials. J Am Acad Dermatol. 2011;64(6):1035-50.

31. Härmark L, Raine J, Leufkens H, Edwards IR, Moretti U, Sarinic VM, et al. Patient-Reported Safety Information: A Renaissance of Pharmacovigilance? Drug Saf. 2016;39(10):883-90.

32. van Hunsel F, Härmark L, Rolfes L. Fifteen years of patient reporting -what have we learned and where are we heading to? Expert Opin Drug Saf. 2019;18(6):477-84.

33. Le S, Shafer PO, Bartfeld E, Fisher RS. An online diary for tracking epilepsy. Epilepsy Behav. 2011;22(4):705-9.

34. Rijsbergen M, Niemeyer-van der Kolk T, Rijneveld R, Pinckaers J, Meshcheriakov I, Bouwes Bavinck JN, et al. Mobile e-diary application facilitates the monitoring of patient-reported outcomes and a high treatment adherence for clinical trials in dermatology. J Eur Acad Dermatol Venereol. 2020;34(3):633-9.

35. National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE) v5.0. https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_5x7.pdf. Accessed 8/1/2021.

36. MedDRA Maintenance and Support Services Organization. Medical Dictionary for Regulatory Activities. https://www.meddra.org/. Accessed 8/1/2021.

37. European Medicines Agency. Avastin: EPAR - Product Information. https://www.ema.europa.eu/en/documents/product-information/avastin-epar-product-information_en.pdf. Accessed 2/12/2020.

38. Bao P, Kodra A, Tomic-Canic M, Golinko MS, Ehrlich HP, Brem H. The role of vascular endothelial growth factor in wound healing. J Surg Res. 2009;153(2):347-58.

39. Gardam MA, Keystone EC, Menzies R, Manners S, Skamene E, Long R, et al. Anti-tumour necrosis factor agents and tuberculosis risk: mechanisms of action and clinical management. Lancet Infect Dis. 2003;3(3):148-55.

40. Pondé NF, Lambertini M, de Azambuja E. Twenty years of anti-HER2 therapy-associated cardiotoxicity. ESMO Open. 2016;1(4):e000073.

41. European Medicines Agency. Herceptin: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/herceptin-epar-product-information_en.pdf. Accessed 2/12/2020.

42. European Medicines Agency. Perjeta: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/perjeta-epar-product-information_en.pdf. Accessed 2/12/2020.

43. Cohan SL, Lucassen EB, Romba MC, Linch SN. Daclizumab: Mechanisms of Action, Therapeutic Efficacy, Adverse Events and Its Uncovering the Potential Role of Innate Immune System Recruitment as a Treatment Strategy for Relapsing Multiple Sclerosis. Biomedicines. 2019;7(1).

44. Stefansdottir G, Knol MJ, Arnardottir AH, van der Elst ME, Grobbee DE, Leufkens HG, et al. Safety learning from drugs of the same class: room for improvement. Clin Pharmacol Ther. 2012;91(5):872-80.

45. European Medicines Agency. Imfinzi: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/imfinzi-epar-public-assessment-report_en.pdf. Accessed 2/12/2020.

46. European Medicines Agency. Imfinzi: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/imfinzi-epar-product-information_en.pdf. Accessed 2/12/2020.

47. European Medicines Agency. Sarclisa: EPAR - Public assessment report. https://www.ema.europa.eu/en/documents/assessment-report/sarclisa-epar-public-assessment-report_en.pdf. Accessed 1/2/2021.

48. Metcalfe C, Dougall T, Bird C, Rigsby P, Behr-Gross ME, Wadhwa M, et al. The first World Health Organization International Standard for infliximab products: A step towards maintaining harmonized biological activity. MAbs. 2019;11(1):13-25.

49. The European Directorate for the Quality of Medicines & healthcare (EDQM). Development of New Ph. Eur. Bioassay "Horizontal" Standards: An Insight into the Anti-TNF-alpha Product Class Collaborative Study. https://cdn.ymaws.com/www.casss.org/resource/resmgr/cmc_euro/cmc2020_virtual/buda_mihaela_slides.pdf. Accessed 18/12/2020.

50. European Medicines Agency. 2019 Annual Report on EudraVigilance for the European Parliament, the Council and the Commission. https://www.ema.europa.eu/en/documents/report/2019-annual-report-eudravigilance-european-parliament-council-commission-reporting-period-1-january_en.pdf. Accessed 28/12/2020.

51. Potts J, Genov G, Segec A, Raine J, Straus S, Arlett P. Improving the Safety of Medicines in the European Union: From Signals to Action. Clin Pharmacol Ther. 2020;107(3):521-9.

52. Arlett P, Straus S, Rasi G. Pharmacovigilance 2030: Invited Commentary for the January 2020 "Futures" Edition of Clinical Pharmacology and Therapeutics. Clin Pharmacol Ther. 2020;107(1):89-91.

53. Ebbers HC, Mantel-Teeuwisse AK, Moors EH, Sayed Tabatabaei FA, Schellekens H, Leufkens HG. A cohort study exploring determinants of safety-related regulatory actions for biopharmaceuticals. Drug Saf. 2012;35(5):417-27.

54. European Medicines Agency. New study suggests risk of birth defects in babies born to women on HIV medicine dolutegravir. https://www.ema.europa.eu/en/documents/press-release/new-study-suggests-risk-birth-defects-babies-born-women-hiv-medicine-dolutegravir_en.pdf. Accessed 28/12/2020.

55. Zash R, Makhema J, Shapiro RL. Neural-Tube Defects with Dolutegravir Treatment from the Time of Conception. The New England journal of medicine. 2018;379(10):979-81.

56. Schotland P, Racz R, Jackson DB, Soldatos TG, Levin R, Strauss D, et al. Target adverse event profiles for predictive safety in the post-market setting. Clinical Pharmacology & Therapeutics. 2020 Oct 8. doi: 10.1002/cpt.2074. Online ahead of print.

57. Liu M, Wu Y, Chen Y, Sun J, Zhao Z, Chen XW, et al. Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs. J Am Med Inform Assoc. 2012;19(e1):e28-35.

58. Tatonetti NP. The Next Generation of Drug Safety Science: Coupling Detection, Corroboration, and Validation to Discover Novel Drug Effects and Drug-Drug Interactions. Clin Pharmacol Ther. 2018;103(2):177-9.

59. Lorberbaum T, Sampson KJ, Chang JB, Iyer V, Woosley RL, Kass RS, et al. Coupling Data Mining and Laboratory Experiments to Discover Drug Interactions Causing QT Prolongation. Journal of the American College of Cardiology. 2016;68(16):1756-64.

60. Daei A, Soleymani MR, Ashrafi-rizi H, Zargham-Boroujeni A, Kelishadi R. Clinical information seeking behavior of physicians: A systematic review. International Journal of Medical Informatics. 2020;139:104144.

61. Rahmner PB, Eiermann B, Korkmaz S, Gustafsson LL, Gruvén M, Maxwell S, et al. Physicians' reported needs of drug information at point of care in Sweden. British Journal of Clinical Pharmacology. 2012;73(1):115-25.

62. IQVIA Institute for Human Data Science. Medicine Use and Spending in the U.S. A Review of 2017 and Outlook to 2022. https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/medicine-use-and-spending-in-the-us-a-review-of-2017-and-outlook-to-2022.pdf?_=1606734271803. Accessed 2/1/2021.

63. Hendrikx J, Haanen J, Voest EE, Schellens JHM, Huitema ADR, Beijnen JH. Fixed Dosing of Monoclonal Antibodies in Oncology. Oncologist. 2017;22(10):1212-21.

64. European Medicines Agency. Keytruda: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/keytruda-epar-product-information_en.pdf. Accessed 1/12/2020.

65. European Medicines Agency. Opdivo: EPAR - Product information. https://www.ema.europa.eu/en/documents/product-information/opdivo-epar-product-information_en.pdf. Accessed 1/12/2020.

66. Volokhina E, Wijnsma K, van der Molen R, Roeleveld N, van der Velden T, Goertz J, et al. Eculizumab Dosing Regimen in Atypical HUS: Possibilities for Individualized Treatment. Clinical Pharmacology & Therapeutics. 2017;102(4):671-8.

67. Verhoef LM, van den Bemt BJ, van der Maas A, Vriezekolk JE, Hulscher ME, van den Hoogen FH, et al. Down-titration and discontinuation strategies of tumour necrosis factor-blocking agents for rheumatoid arthritis in patients with low disease activity. Cochrane Database Syst Rev. 2019;5(5):Cd010455.

68. Beers E, Egberts TC, Leufkens HG, Jansen PA. Information for adequate prescribing to older patients : an evaluation of the product information of 53 recently approved medicines. Drugs Aging. 2013;30(4):255-62.

69. Rosenfeld RM, Shiffman RN. Clinical practice guideline development manual: a quality-driven approach for translating evidence into action. Otolaryngol Head Neck Surg. 2009;140(6 Suppl 1):S1-43.

70. Montalban X, Gold R, Thompson AJ, Otero-Romero S, Amato MP, Chandraratna D, et al. ECTRIMS/EAN Guideline on the pharmacological treatment of people with multiple sclerosis. Mult Scler. 2018;24(2):96-120.

71. Cochrane. Clinical Study Reports within Cochrane Reviews Consultation Meeting: minutes. https://community.cochrane.org/news/using-clinical-study-reports-data-source-cochrane-reviews-consultation-meeting-report-and-next-steps. Accessed 04/1/2021.

72. Ioannidis JP, Evans SJ, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med. 2004;141(10):781-8.

6

SUMMARY & SAMENVATTING

6.1

SUMMARY

INTRODUCTION

Prior to marketing approval of drugs, extensive research should be performed by the pharmaceutical companies and submitted to the regulatory authorities to ensure the pharmaceutical quality, efficacy, and safety. If, based on the data provided, the regulatory authorities consider that a consistent quality is demonstrated and the benefits of a drug outweigh the risks in the treated population, the drug is approved. However, uncertainties about the safety and efficacy of a drug always remain at the time of marketing approval given the limitations of clinical trials (e.g., limited number and highly selected patients and relatively short follow-up time). Especially the safety profile should be further characterized in the post-marketing phase, for which pharmacovigilance has been put in place. Today's pharmacovigilance system is a proactive system and features different tools to further characterize the safety profile of drugs. These tools include collecting and monitoring of spontaneously reported suspected adverse drug reactions (ADRs) and performing post-marketing safety studies that can make use of a variety of study designs and systems.

As described in *Chapter 1*, for biologicals, specific challenges have been addressed for post-marketing safety learning. These include the difference in nature of ADRs from those known for small molecules, the complexity of the mechanism of action (including interference with the immune system), and difficulties in classifying ADRs according to the established system. Also, the detection of ADRs for biologicals is hampered when the symptoms of the ADRs mimic those of the treated disease. Previous research - including several PhD projects from our group - have addressed post-marketing safety learning for biologicals and studied a variety of safety-related regulatory actions and regulatory activities. The field is continuously evolving and the introduction of biologicals with new mechanisms of action brings additional challenges with them. Therefore, in this thesis, we aimed to provide further insights in post-marketing safety learning for biologicals with a specific focus on the characterization of specific safety issues, dynamics in post-marketing safety learning, and safety information from regulatory and clinical sources.

CHARACTERIZATION OF SPECIFIC SAFETY ISSUES

In *Chapter 2,* we focused on the characterization of specific safety issues for biologicals. In *Chapter 2.1*, we showed that there is a potential association between depression and suicidal ideation and behavior and the use of monoclonal antibodies through their influence on the immune system. Our study assessed spontaneously reported suspected ADRs for 44 monoclonal antibodies that had been authorized in the European Union and/or United States as of 2014. For these monoclonal antibodies, we identified the reports of depression (n = 9455) and suicidal ideation and behavior (n = 1770) in the WHO global database of individual case safety reports, VigiBase. The strongest association was found for natalizumab and belimumab as compared to bevacizumab, both for depression (reporting odds ratio (ROR) 5.7, 95% confidence interval [CI] 5.0 – 6.4; and ROR 5.1, 95% CI 4.2 – 6.2) and suicidal ideation and behavior (ROR 12.0, 95% CI 7.9 – 18.3; and ROR 20.2, 95% CI 12.4 – 33.0). When grouping the monoclonal antibodies according to their influence on the immune system, those suppressing the immune system showed a relatively higher reporting frequency, i.e. ROR 1.9 (95% CI 1.8 – 2.0) for depression and ROR 3.6 (95% CI 3.0 – 4.4) for suicidal ideation and behavior as compared to

the monoclonal antibodies not directly targeting the immune system. Whereas we focused in *Chapter 2.1* on a potential new safety issue, we further characterized a known safety issue in *Chapter 2.2*. We used data from clinical practice to study the incidence, longitudinal pattern, and potential risk factors of thyroid disorders in a cohort of patients treated with PD-1/PD-L1 inhibitors. Of the 465 patients who were included in this study, 13% (n = 58) developed thyroid disorders. Of these 58 patients that developed thyroid disorders, 19% (n = 11) had isolated hypothyroidism, which occurred after a median of 69 days. The remaining 81% (n = 47) of the patients developed hyperthyroidism, which occurred, if isolated, after a median of 55 days (48%, n = 28). Hyperthyroidism occurred after 21 days for those patients who subsequently developed hypothyroidism at a median of 48 days later (33%, n = 19). The thyroid levels normalized within a median of 55 days. A specific pattern with a rapid decline in thyroid-stimulating hormone values after initiation of the therapy followed by an increase was observed in about 4% (n = 14) of the patients. This pattern was more prominently observed for female melanoma patients treated with the combination of a PD-1/PD-L1 inhibitor and ipilimumab. Female patients and patients with pre-existing thyroid disorders were at increased risk of developing thyroid disorders (odds ratio [OR]: 2.04 [95% CI: 1.14 – 3.70], and OR: 4.31 [95% CI: 1.47 – 12.61], respectively).

6.1 DYNAMICS IN POST-MARKETING SAFETY LEARNING

In *Chapter 3,* we studied different dynamics in post-marketing safety learning. In *Chapter 3.1*, we showed that, in contrast to what has been reported for small molecules, the dosing information for biologicals is rarely reduced for safety reasons. Our study assessed post-marketing learning regarding dosing information for the 71 biologicals that were authorized by the European Medicines Agency between 2007 and 2014 that were followed up until December 2016. For these 71 biologicals, the dosing information for the initial indication was changed during follow-up for eight biologicals (11%), which was rarely reduced for safety reasons. For 30 products (42%), the indication was extended at least once. We did not observe changes in dosing information for the extended indications (n = 59) during follow-up. In *Chapter 3.2,* we concluded that there is room for improvement regarding post-marketing safety learning among drugs with commonalities in the mechanism of action. We showed that the ADRs described in the summary of product characteristics (SmPC) of the TNF-α inhibitors differ substantially. Our study assessed the overlap in ADRs described in the SmPC of the TNF-α inhibitors throughout the drug's life cycle. At the end of follow-up (31 December 2019), 293 unique ADRs were described in the SmPCs of the five included TNF-α inhibitors. Of the 293 ADRs, 133 (45%) were described in the SmPC of one TNF-α inhibitor and 39 (13%) in the SmPCs of all five TNF-α inhibitors. Serious ADRs and ADRs classified as important risks by the regulators were described approximately four times more often in a second SmPC than ADRs not classified as such.

SAFETY INFORMATION FROM REGULATORY AND CLINICAL SOURCES

In *Chapter 4,* we evaluated and compared the safety information provided in regulatory and clinical sources. In the study described in *Chapter 4.1*, we showed that different methods were used to

identify and classify thyroid disorders associated with the PD-1/PD-L1 inhibitors in the clinical trials and observational studies. The 38 included clinical trials reported a non-specific method to identify patients as having thyroid disorders since no reference ranges for the thyroid hormones were provided. The observational studies were more specific, with the majority (n = 23, 82%) specifying the reference ranges for the thyroid hormones. Multiple subclassifications of thyroid disorders (e.g., hyperthyroidism, hypothyroidism) were reported, with the reporting of subclinical thyroid disorders in the observational studies as main difference in reporting between the clinical trials and observational studies. Moreover, in the clinical trials a specific assessment was performed to subclassify thyroid disorders as being immune-related, for which the assessment differed among the different PD-1/PD-L1 inhibitors and, for some, between the clinical trials for one drug. In *Chapter 4.2,* we showed that there is substantial discordance in the reporting of adverse events on the same phase 3 trials between regulatory and clinical information sources. Our study assessed the adverse events of eight biologicals used for the treatment of multiple sclerosis for which a total of 707 adverse events were reported in the European public assessment report (EPAR) and/or scientific publication. Approximately one-third of the adverse events was reported in both the EPAR and scientific publication, one-third was only reported in the EPAR and one-third only in the scientific publication. Serious adverse events and adverse events that regulators classified as 'important identified risk' were significantly more often reported in both sources compared to adverse events not classified as such (respectively, 38% vs. 30% and 49% vs. 30%).

IMPLICATIONS

In *Chapter 5*, we have put the findings of the studies presented in this thesis in broader perspective. First, we discussed the dynamics in the indication of use and dosing information for post-marketing learning and concluded that dynamics in dosing information play an important role in the context of post-marketing learning and could be optimized for risk minimization purposes and patient's convenience. Second, we described the challenges related to post-marketing safety learning regarding the detection and classification of ADRs. We concluded that the detection of ADRs for biologicals could be facilitated by intensifying the role of the patients and implementing monitor recommendations to be applied in clinical practice. For the classification of ADRs, efforts should be made to standardize them, especially those that can be objectively detected by using biomarkers. Third, we discussed that post-marketing safety learning could be optimized through cross-learning among drugs with commonalities in the mechanism of action. For this, drugs with commonalities in the mechanism of actions should be grouped according to these commonalities and existing methods could be complemented by using prediction models that combine data from multiple information sources. At last, we discussed the gap between safety information from regulatory and clinical sources and concluded that this gap can be closed by improving the clinical applicability of regulatory information sources and increasingly incorporating information from clinical practice into the regulatory system. With these recommendations, post-marketing safety learning for biologicals can be facilitated from a regulatory and clinical perspective.

6.2

SAMENVATTING

INTRODUCTIE

Voordat een geneesmiddel op de markt wordt toegelaten, dient uitgebreid onderzoek te worden gedaan naar de farmaceutische kwaliteit, werkzaamheid en veiligheid van het geneesmiddel. Deze studies worden uitgevoerd door farmaceutische bedrijven en daarna aan de regulatoire autoriteiten voorgelegd ter beoordeling. Als de regulatoire autoriteiten, op basis van de verstrekte gegevens, van mening zijn dat een consistente kwaliteit is aangetoond en dat de voordelen van het geneesmiddel opwegen tegen de risico's, wordt het geneesmiddel goedgekeurd. Onzekerheden over de veiligheid en werkzaamheid van het geneesmiddel blijven echter altijd bestaan op het moment van markttoelating, gezien de beperkingen van de uitgevoerde klinische studies (bijv. het aantal geïncludeerde patiënten is beperkt, er heeft een strenge selectie van de patiënten plaatsgevonden en de follow-up duur is relatief kort). Vooral het veiligheidsprofiel dient verder in kaart te worden gebracht na markttoelating, waarvoor het geneesmiddelenbewakingssysteem is ingericht. Het huidige geneesmiddelenbewakingssysteem is een proactief systeem en beschikt over verschillende instrumenten om het veiligheidsprofiel van geneesmiddelen verder in kaart te brengen. Deze instrumenten omvatten het verzamelen en monitoren van spontane meldingen van vermoede bijwerkingen en het uitvoeren van veiligheidsonderzoeken na markttoelating, waarbij gebruik kan worden gemaakt van verschillende onderzoeksopzetten en -systemen.

In *Hoofdstuk 1* beschrijven we dat er voor biologische geneesmiddelen specifieke uitdagingen zijn voor de verwerving van kennis over het veiligheidsprofiel nadat het geneesmiddel op de markt is toegelaten. De uitdagingen omvatten onder meer het verschil in aard van bijwerkingen ten opzichte van welke bekend zijn voor kleine moleculen, de complexiteit van het werkingsmechanisme (inclusief interactie met het immuunsysteem), moeilijkheden bij het classificeren van bijwerkingen volgens de vastgestelde systemen en het feit dat de detectie van bijwerkingen belemmerd wordt als de symptomen van de bijwerkingen overeenkomen met die van de behandelde ziekte. Er is eerder onderzoek gedaan naar de verwerving van kennis over het veiligheidsprofiel van biologische geneesmiddelen na markttoelating, waarin een verscheidenheid aan veiligheidsgerelateerde regulatoire maatregelen en regulatoire activiteiten is bestudeerd. Het veld is echter voortdurend in ontwikkeling en de introductie van biologische geneesmiddelen met nieuwe werkingsmechanismen zorgt voor extra uitdagingen. Daarom hebben wij verdere inzichten verschaft in het verwerven van kennis over het veiligheidsprofiel van biologische geneesmiddelen na markttoelating. Hierbij hebben wij ons specifiek gericht op het in kaart brengen van specifieke veiligheidskwesties, de dynamiek in de verwerving van kennis over het veiligheidsprofiel na markttoelating en veiligheidsinformatie beschreven in informatiebronnen van regulatoire autoriteiten en de klinische praktijk.

KARAKTERISERING VAN SPECIFIEKE VEILIGHEIDSKWESTIES

In *Hoofdstuk 2* richtten wij ons op het in kaart brengen van specifieke veiligheidskwesties. In *Hoofdstuk 2.1* toonden we aan dat er een mogelijk verband is tussen depressie, gedachte aan zelfmoord en suïcidaal gedrag en het gebruik van monoklonale antilichamen middels hun invloed op het immuunsysteem. In deze studie bestudeerden we spontane meldingen van vermoede bijwerkingen voor de 44 monoklonale antilichamen die in de Europese Unie en/of de Verenigde

Staten waren goedgekeurd tot 2014. Voor deze monoklonale antilichamen hebben we de meldingen van depressie (n = 9455) en gedachte aan zelfmoord en suïcidaal gedrag (n = 1770) geïdentificeerd in VigiBase. De associatie was het sterkst voor natalizumab en belimumab in vergelijking met bevacizumab, zowel voor depressie (reporting odds ratio (ROR) 5,7, 95% betrouwbaarheidsinterval [BI] 5,0 – 6,4; en ROR 5,1, 95% BI 4,2 – 6,2) als voor de gedachte aan zelfmoord en suïcidaal gedrag (ROR 12,0, 95% BI 7,9 – 18,3; en ROR 20,2, 95% BI 12,4 – 33,0). Wanneer de monoklonale antilichamen gegroepeerd werden op basis van hun invloed op het immuunsysteem, vertoonden degenen die het immuunsysteem onderdrukken een hogere ROR, d.w.z. 1,9 (95% BI 1,8 – 2,0) voor depressie en 3,6 (95% BI 3,0 – 4,4) voor de gedachte aan zelfmoord en suïcidaal gedrag in vergelijking met de monoklonale antilichamen die niet rechtstreeks van invloed zijn op het immuunsysteem. Waar we ons in *Hoofdstuk 2.1* concentreerden op het bestuderen van een potentieel nieuw veiligheidsprobleem, hebben we een bekend veiligheidsprobleem verder gekarakteriseerd in *Hoofdstuk 2.2*. Voor deze studie hebben we gebruik gemaakt van data uit de klinische praktijk om de incidentie, het longitudinale patroon en de mogelijke risicofactoren van schildklieraandoeningen te bestuderen in een cohort van patiënten behandeld met PD-1/PD-L1-remmers. Van de 465 patiënten die in deze studie werden geïncludeerd, ontwikkelde 13% (n = 58) schildklieraandoeningen. Geïsoleerde hypothyreoïdie werd waargenomen bij 19% (n = 11) van de patiënten die schildklieraandoeningen ontwikkelden en trad op na een mediaan van 69 dagen. De overige 81% (n = 47) van de patiënten ontwikkelde hyperthyreoïdie, die, indien geïsoleerd, optrad na een mediaan van 55 dagen (48%, n = 28). Hyperthyreoïdie trad op na 21 dagen voor degenen die vervolgens hypothyreoïdie ontwikkelden met een mediaan van 48 dagen later (33%, n = 19). De schildklierwaarden normaliseerden binnen een mediaan van 55 dagen. Een specifiek patroon met een snelle afname van de waarden van het schildklierstimulerend hormoon na aanvang van de therapie gevolgd door een toename hiervan werd waargenomen bij ongeveer 4% (n = 14) van de patiënten. Dit patroon werd met name waargenomen bij vrouwelijke melanoompatiënten die werden behandeld met de combinatie van een PD-1/PD-L1-remmer en ipilimumab. Vrouwelijke patiënten en patiënten met reeds bestaande schildklieraandoeningen hadden een verhoogd risico op het ontwikkelen van schildklieraandoeningen tijdens de behandeling met PD-1/PD-L1-remmers (odds ratio [OR]: 2,04 [95% BI: 1,14 – 3,70], en OR: 4,31 [95% BI: 1,47 – 12,61], respectievelijk).

DYNAMIEK IN HET VERWERVEN VAN KENNIS OVER HET VEILIGHEIDSPROFIEL NA MARKTTOELATING

In *Hoofdstuk 3* bestudeerden we verschillende dynamieken in het verwerven van kennis over het veiligheidsprofiel nadat geneesmiddelen zijn toegelaten op de markt. In *Hoofdstuk 3.1* toonden wij aan dat de doseringsinformatie voor biologische geneesmiddelen, in tegenstelling tot wat gerapporteerd is voor kleine moleculen, zelden wordt verlaagd om veiligheidsredenen. In onze studie bestudeerden wij de dynamiek in het verwerven van kennis met betrekking tot de doseringsinformatie van biologische geneesmiddelen na markttoelating voor de 71 biologische geneesmiddelen die tussen 2007 en 2014 waren goedgekeurd door het Europees Geneesmiddelenbureau en volgden ze tot december 2016. Van deze 71 biologische geneesmiddelen werd de doseringsinformatie voor

de initiële indicatie gewijzigd voor acht geneesmiddelen (11%) tijdens de follow-up, welke zelden verlaagd werd om veiligheidsredenen. Voor 30 producten (42%) is de indicatie minimaal één keer uitgebreid. We namen geen veranderingen in de doseringsinformatie waar voor de indicatie-uitbreidingen (n = 59) tijdens de follow-up. In *Hoofdstuk 3.2* hebben wij geconcludeerd dat er ruimte is voor voorbetering met betrekking tot het verwerven van kennis over het veiligheidsprofiel tussen geneesmiddelen met overeenkomsten in hun werkingsmechanisme. Wij lieten zien dat de bijwerkingen die worden beschreven in de samenvatting van de productkenmerken (SmPC) van geneesmiddelen met hetzelfde werkingsmechanisme aanzienlijk verschillen. In deze studie bestudeerden wij de overlap in bijwerkingen beschreven in de SmPC van de TNF-α-remmers gedurende hun levenscyclus. Aan het einde van de follow-up (31 december 2019) werden 293 unieke bijwerkingen beschreven in de SmPC's van de vijf geïncludeerde TNF-α-remmers. Van de 293 bijwerkingen werden er 133 (45%) beschreven in de SmPC van één TNF-α-remmer en 39 (13%) in de SmPC's van alle vijf TNF-α-remmers. Ernstige bijwerkingen en bijwerkingen geclassificeerd als belangrijke risico's door de regulatoire autoriteiten werden ongeveer vier keer vaker beschreven in een tweede SmPC dan bijwerkingen die niet als zodanig waren geclassificeerd.

VEILIGHEIDSINFORMATIE UIT BRONNEN VAN REGULATOIRE AUTORITEITEN EN DE KLINISCHE PRAKTIJK

In *Hoofdstuk 4* hebben we de veiligheidsinformatie uit bronnen van regulatoire autoriteiten en de klinische praktijk geëvalueerd en vergeleken. In de studie beschreven in *Hoofdstuk 4.1* hebben we aangetoond dat verschillende methoden worden gebruikt om schildklieraandoeningen die geassocieerd zijn met de PD-1/PD-L1-remmers te identificeren en classificeren in klinische studies en observationele studies. In alle 38 klinische studies was de omschreven methode die werd gebruikt om patiënten te identificeren met schildklieraandoeningen niet specifiek, aangezien er geen referentiewaarden voor de schildklierhormonen werden beschreven. Deze methode was specifieker in de observationele studies, aangezien in de meeste observationele studies (n = 23, 82%) de referentiewaarden voor de schildklierhormonen beschreven werden. Meerdere classificaties van schildklieraandoeningen (bijv. hyperthyreoïdie, hypothyreoïdie) werden gerapporteerd met de rapportage van subklinische schildklieraandoeningen in de observationele studies als belangrijkste verschil in rapportage tussen klinische en observationele studies. Bovendien werd in de klinische studies een specifieke beoordeling uitgevoerd om schildklieraandoeningen te classificeren als immuun-gerelateerd, waarvoor de beoordeling verschilde tussen de verschillende PD-1/PD-L1-remmers en, voor een aantal PD-1/PD-L1-remmers, tussen de klinische studies voor één PD-1/PD-L1-remmer. In *Hoofdstuk 4.2* toonden we aan dat er aanzienlijke verschillen bestaan in de rapportage van bijwerkingen op basis van dezelfde fase 3 studies tussen bronnen van regulatoire autoriteiten en de klinische praktijk. In onze studie bestudeerden we acht biologische geneesmiddelen die worden gebruikt voor de behandeling van multiple sclerose, waarvoor het Europese publieke beoordelingsrapport (EPAR) en/of de wetenschappelijke publicatie in totaal 707 bijwerkingen rapporteerden. Ongeveer een derde van de bijwerkingen werd gerapporteerd in zowel de EPAR als in de wetenschappelijke publicatie, een derde werd alleen in de EPAR gerapporteerd en een derde

alleen in de wetenschappelijke publicatie. Ernstige bijwerkingen en bijwerkingen die regulatoire autoriteiten classificeerden als 'belangrijk geïdentificeerd risico' werden significant vaker gerapporteerd in beide informatiebronnen in vergelijking met bijwerkingen die niet als zodanig waren geclassificeerd (respectievelijk 38% versus 30% en 49% versus 30%).

IMPLICATIES

In *Hoofdstuk 5* hebben we de bevindingen van de studies gepresenteerd in dit proefschrift in een breder perspectief geplaatst. Als eerste beschreven wij de veranderingen in zowel de indicatie waarvoor het geneesmiddel wordt gebruikt als doseringsinformatie voor het verwerven van kennis na markttoelating. We concludeerden dat de dynamiek in doseringsinformatie een belangrijke rol speelt in de context van het verwerven van kennis na markttoelating en geoptimaliseerd kan worden uit risicominimalisatie overwegingen en het verbeteren van de therapie voor de patiënt. Als tweede bediscussieerden wij de uitdagingen gerelateerd aan het identificeren en classificeren van bijwerkingen. Hiervoor concludeerden wij dat de detectie van bijwerkingen voor biologische geneesmiddelen zou kunnen worden verbeterd door de rol van de patiënten te intensiveren en concrete aanbevelingen voor monitoring te implementeren die in de klinische praktijk moeten worden toegepast. Voor de classificatie van bijwerkingen moeten inspanningen worden geleverd om ze te standaardiseren, vooral die bijwerkingen welke objectief kunnen worden gedetecteerd met behulp van biomarkers. Als derde beschreven wij dat het verwerven van kennis over het veiligheidsprofiel na markttoelating geoptimaliseerd kan worden door kennis uit te wisselen tussen geneesmiddelen met overeenkomsten in het werkingsmechanisme, waarbij de geneesmiddelen gegroepeerd kunnen worden op basis van deze overeenkomsten. Daarnaast kunnen bestaande methoden voor het verwerven van kennis na markttoelating gecombineerd worden met predictiemodellen die gebruik maken van gegevens uit meerdere informatiebronnen. Als laatste bediscussieerden wij het verminderen van het verschil tussen veiligheidsinformatie uit bronnen van regulatoire autoriteiten en de klinische praktijk. Het verschil tussen veiligheidsinformatie beschreven in bronnen van regulatoire autoriteiten en de klinische praktijk kan worden verminderd door de klinische toepasbaarheid van informatiebronnen van regulatoire autoriteiten te verbeteren en door meer informatie uit de klinische praktijk in het regulatoire systeem op te nemen. Middels deze aanbevelingen kan het verwerven van kennis over het veiligheidsprofiel na markttoelating verbeterd worden vanuit een regulatoir en klinisch perspectief.

7

APPENDICES

7.1

DANKWOORD

Door de hulp van velen heb ik dit promotietraject kunnen afronden. Via deze weg wil ik iedereen bedanken die een, zowel wetenschappelijke als sociale, bijdrage heeft geleverd aan dit proefschrift. Graag wil ik een aantal van hen in het bijzonder bedanken.

Allereerst wil ik mijn promotieteam bestaande uit prof. dr. Bert Leufkens, prof. dr. Toine Egberts, dr. Helga Gardarsdottir en dr. Thijs Giezen bedanken. Bert, jij kent de regulatoire wereld als geen ander, maar verliest ook nooit de klinische kant uit het oog. Ook heb ik veel van je geleerd door de data op verschillende manieren te bekijken, wat vaak nieuwe inzichten opleverde. Toine, jij hebt me gestimuleerd (soms met wat tegenzin van mijn kant) om uit mijn comfortzone te stappen, wat mij verder heeft gebracht. Je hield het tempo in de studies door inzichtelijk te maken wat er nodig was en deze hulp te bieden. Helga, als dagelijkse begeleider op de universiteit kon en mocht ik altijd even bij je binnenlopen. Met een praktische oplossing zorgde je ervoor dat ik dan weer verder kon. Ook jouw epidemiologische kennis en verhelderende suggesties waren onmisbaar. Thijs, jouw kennis over dit onderwerp en ideeën hebben richting gegeven aan verschillende studies. Het was heel fijn dat jij dit traject zelf had doorlopen en ik ook bij jou altijd terecht kon.

Aukje, ook al maak je aan het eind van het traject officieel geen deel meer uit van het promotieteam, jij hebt er in het eerste jaar voor gezorgd dat ik een goede start kon maken. Hierbij waren jouw ervaringen met eerdere promotietrajecten over dit onderwerp van belangrijke toegevoegde waarde.

The members of the reading committee, prof. dr. Marieke de Bruin, dr. Marjon Pasmooij, prof. dr. Eugène van Puijenbroek, prof. dr. Gabe Sonke, prof. dr. Gianluca Trifiró, I would like to thank for their assessment of this thesis.

Graag wil ik ook de verschillende coauteurs bedanken voor hun bijdrage. Jarno, bedankt dat ik gebruik heb mogen maken van jouw tool om de bijwerkingen in kaart te brengen en voor de scherpe input voor het manuscript. Patrick, dankjewel voor de geboden mogelijkheid om gebruik te maken van de data uit VigiBase en de hulp bij de totstandkoming van het manuscript. Voor de studie in hoofdstuk 2.2 heb ik gebruik mogen maken van de data uit UPOD. Hiervoor wil ik Cornelia Hulsbergen-Veelken en Mark de Groot graag bedanken voor de extractie van de data en hun kennis over de beschikbare data. Voor deze studie was ook de klinische input van Karijn Suijkerbuijk onmisbaar, dank daarvoor.

De vele collega's op de universiteit hebben er voor gezorgd dat ik de afgelopen jaren een leerzame maar zeker ook leuke tijd heb gehad.

Lourens, het was heel fijn om dit traject samen met jou te doorlopen en de uitdagingen van het combinatietraject met elkaar te kunnen delen. Daarnaast waren de verschillende tripjes naar Kopenhagen en Genève ook heel gezellig. Rick, ik was heel blij dat jij een aantal maanden na mijn

start ook op de universiteit kwam werken. We mochten maar kort naast elkaar zitten maar gelukkig bleken de lunchwandelingen een goede vervanging. Joris, de analyses voor de verschillende studies had ik nooit kunnen uitvoeren zonder jouw hulp met R. Gelukkig kon ik jou bijstaan met regeldingen en ik heb je vrijdagmiddagafleiding het laatste jaar zeker gemist. Joost, de dagen op de universiteit waren met jou nooit saai, maar je was ook altijd vol belangstelling en zorgde voor wat motiverende woorden als dat nodig was. Daarnaast heb je mij over veel verschillende (heel nuttige) onderwerpen wat geleerd en is mijn woordenschat er op vooruit gegaan. Pieter, ook jij hebt mij veel bijgebracht, de challenges heb ik helaas alleen nooit gehaald. Heel veel succes met het afronden van jouw proefschrift. Delphi, we hebben heel wat uur naast elkaar gewerkt, maar waren ook kamergenoten tijdens de skitrips, dankjewel voor de adviezen en gezelligheid. Renske, het was heel fijn om de laatste loodjes samen te kunnen delen. Richelle, jouw vrolijkheid en enthousiasme zorgden ervoor dat de sfeer er altijd goed in zat. Marianne, dankjewel dat ik altijd bij jou terecht kon met vragen over coördinatietaken. Ineke, Suzanne, Anja en Paula, bedankt voor al jullie hulp en de gezellige gesprekken in de ochtend.

Alle mede PhD studenten, Ali, Amos, Debbie, Doerine, Gert-Jan, Hedy, Iris, Jeroen, Jet, Joris, Lenneke, Li, Mariette, Marle, Melissa, Milou, Mirjam, Nick, Rachel, Rosanne, Sander en andere collega's van de afdeling wil ik graag bedanken voor de gezelligheid tijdens alle uitjes, skitrips, borrels en natuurlijk frituurvrijdag.

7.1

Naast mijn collega's van de universiteit wil ik graag mijn CBG collega's bedanken voor hun interesse en flexibiliteit. Ik ben heel blij dat we collega's blijven. Joan, dankjewel voor de kansen die je me hebt geboden en het vertrouwen om mij verder te kunnen ontwikkelen bij het CBG.

Heshu, tijdens mijn onderzoeksproject onder jouw begeleiding is mijn interesse voor het onderzoek ontstaan. Ik ben blij dat we nog steeds contact hebben en we regelmatig bijkletsen.

Ook mijn vrienden en familie wil ik graag bedanken voor hun interesse in mijn onderzoek en de ontspanning tijdens de afgelopen jaren.

Farmacievrienden, na onze studie farmacie zijn we veel verschillende richtingen op gegaan. Gelukkig blijven de farmagrappen en hebben jullie de afgelopen jaren voor veel gezellige afleiding gezorgd, dankjulliewel daarvoor. Robbert, het samen studeren voor de wiskunde examens op de middelbare school en later de farmacietentamens liggen ver achter ons, maar ik ben heel blij dat we onze gedeelde hobby van avondjes slechte televisieprogramma's kijken nog steeds kunnen delen. Willem, tijdens de thuiswerkperiode werden we een soort van collega's en het was heel fijn om af en toe wat andere muren om me heen te kunnen hebben.

Leukste vriendinnengroep uit Nijmegen, het is alweer heel wat jaar geleden dat we samen eindexamen deden, maar gelukkig zien we elkaar nog steeds en voelt het altijd vertrouwd. Hopelijk kunnen we komend jaar de tradities van de kerstdiners en vierdaagse weer voortzetten.

Sterre, ontelbaar veel sportlesjes hebben wij de afgelopen jaren samen gevolgd. Het was heel fijn om alleen te denken aan wanneer de mountain climbers voorbij zijn.

Bar, Bel, Lies en Meer, zeker de laatste periode van mijn promotietraject hebben jullie mij er doorheen geholpen en stonden jullie altijd voor mij klaar. Dat er nog maar heel veel gezellige etentjes, sportsessies en vakanties mogen volgen. Zeker ook met Duco, Ray en Tim.

Sil, onze werelden staan ver van elkaar af maar je was altijd geïnteresseerd in mijn onderzoek en gelukkig weet je inmiddels ook dat dit boekje niet over A. Vogel gaat. Papa en mama, jullie hebben mij altijd gesteund en geholpen waar dat kon. Bedankt voor alles!

7.2

LIST OF CO-AUTHORS

A.C.G. (Toine) Egberts

Division of Pharmacoepidemiology & Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Utrecht, The Netherlands
Department of Clinical Pharmacy, Division Laboratory, Pharmacy and Biomedical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands

H. (Helga) Gardarsdottir

Division of Pharmacoepidemiology & Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Utrecht, The Netherlands
Department of Clinical Pharmacy, Division Laboratory, Pharmacy and Biomedical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands
Department of Pharmaceutical Sciences, School of Health Sciences, University of Iceland, Reykjavik, Iceland

T.J. (Thijs) Giezen

Foundation Pharmacy for Hospitals in Haarlem, Haarlem, The Netherlands
Department of Clinical Pharmacy, Spaarne Gasthuis, Haarlem/Hoofddorp, The Netherlands
Medicines Evaluation Board, Utrecht, The Netherlands

M.C.H. (Mark) de Groot

Central Diagnostic Laboratory, Division Laboratory, Pharmacy and Biomedical Genetics, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

J. (Jarno) Hoekman

Division of Pharmacoepidemiology & Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Utrecht, The Netherlands
Innovation Studies, Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, The Netherlands

C.A.R. (Cornelia) Hulsbergen-Veelken

Central Diagnostic Laboratory, Division Laboratory, Pharmacy and Biomedical Genetics, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

H.G.M. (Bert) Leufkens

Division of Pharmacoepidemiology & Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Utrecht, The Netherlands

A.K. (Aukje) Mantel-Teeuwisse

Division of Pharmacoepidemiology & Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Utrecht, The Netherlands

P.C. (Patrick) Souverein

Division of Pharmacoepidemiology & Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Utrecht, The Netherlands

K.P.M. (Karijn) Suijkerbuijk

Department of Medical Oncology, Cancer Center, University Medical Center Utrecht, Utrecht, The Netherlands

7.3

LIST OF PUBLICATIONS

PUBLICATIONS RELATED TO THIS THESIS

Minnema LA, Giezen TJ, Hoekman J, Egberts TCG, Leufkens HGM, Gardarsdottir H.
Regulatory Safety Learning Driven by the Mechanism of Action: The Case of TNF- Inhibitors.
Clin Pharmacol Ther 2020. Online ahead of print.
https://pubmed.ncbi.nlm.nih.gov/33278830/

Minnema LA, Giezen TJ, Egberts TCG, Leufkens HGM, Gardarsdottir H.
Adverse events related to biologicals used for patients with multiple sclerosis: a comparison between information originating from regulators and information originating from the scientific community.
Eur J Neurol 2020;27(7):1250-6.
https://pubmed.ncbi.nlm.nih.gov/32298524/

Minnema LA, Giezen TJ, Souverein PC, Egberts TCG, Leufkens HGM, Gardarsdottir H.
Exploring the Association between Monoclonal Antibodies and Depression and Suicidal Ideation and Behavior: A VigiBase Study.
Drug Saf 2019;42(7):887-95.
https://pubmed.ncbi.nlm.nih.gov/30617497/

Minnema LA, Giezen TJ, Gardarsdottir H, Egberts TCG, Leufkens HGM, Mantel-Teeuwisse AK.
Post-marketing dosing changes in the label of biologicals.
Br J Clin Pharmacol 2019;85(4):715-21.
https://pubmed.ncbi.nlm.nih.gov/30593087/

PUBLICATION NOT RELATED TO THIS THESIS

Abdullah-Koolmees H, Gardarsdottir H, Minnema LA, Elmi K, Stoker LJ, Vuyk J, Goedhart LE, Egberts TCG, Heerdink ER.
Predicting rehospitalization in patients treated with antipsychotics: a prospective observational study.
Ther Adv Psychopharmacol 2018;8(8):213-29.
https://pubmed.ncbi.nlm.nih.gov/30065813/

7.4

ABOUT THE AUTHOR

Lotte Minnema was born in 1991 in Nijmegen, the Netherlands. She graduated at the Stedelijk Gymnasium Nijmegen in 2010. Hereafter, she started studying at Utrecht University and obtained her Bachelor's degree in 2013, followed by her Master's degree in Pharmacy in 2016.

Following graduation, Lotte started her PhD project at the Division of Pharmacoepidemiology and Clinical Pharmacology of the Utrecht Institute for Pharmaceutical Sciences (UIPS), Faculty of Science, Utrecht University. The PhD project was conducted under the umbrella of the Regulatory Science collaboration between the Dutch Medicines Evaluation Board and UIPS. Lotte combined her PhD project with a position as pharmacovigilance assessor at the Dutch Medicines Evaluation Board. As part of her doctoral education, Lotte followed a training programme that included the postgraduate Master's programme in Epidemiology (Master's degree obtained in 2020) and courses on general research skills and academic competencies. Furthermore, she contributed to educational activities of the pharmacy curriculum at Utrecht University, including coordinating the Pharmacy Master's course 'Oncology and pain', and organized two international symposiums for young researchers in regulatory science.

Lotte currently works as a pharmacovigilance assessor at the Dutch Medicines Evaluation Board.