DATE: OCTOBER 1st, 2019

Master's Thesis – master Innovation Sciences

THE INFLUENCE OF ORGANIZATIONAL REPUTATION ON DRUG REVIEW TIMES: A QUANTITATIVE ANALYSIS OF NEW DRUG APPROVALS AT THE EUROPEAN MEDICINES AGENCY.

STUDENT	Remo Schimmel
STUDENT NUMBER	4135636
E-MAIL	r.schimmel@students.uu.nl
TEL.	+31 6 81881980
SUPERVISOR	Dr. Jarno Hoekman
E-MAIL	j.hoekman@uu.nl

Utrecht University

Abstract

The theory of social judgements of organizations received considerable attention in organizational literature during the last decade. Studies focused on demarcating organizational legitimacy, reputation and status and how these affect decision-making. Judgements about organizational reputation are used as heuristic to estimate future product quality based on previous interactions with the organization holding the reputation. In management literature, organizational reputation is threefold, divided into familiarity (being known), expectations about future outcomes based on prior performance (being known for something) and general impressions about the organization (generalized favorability). Studies proved that when uncertainty exists about product outcome, judgements about the organizational reputation are used to aid decision-making. This is especially relevant for industries were product quality cannot be assessed before large-scale usage. The pharmaceutical industry matches these conditions and is researched extensively, primarily focusing on the case of the US Food and Drug Administration (FDA). These studies proved that being known and generalized favorability facilitate decision-making, but largely neglected the effects of being known for something.

This study aims to provide insights in the relationship between three types of organizational reputation and drug review times, by empirically testing them in the context of new drug approvals in Europe. This allowed replication of the variables used in research on the FDA, while extending to newly introduced indicators that measured the applicants prior performance regarding quality issues and product discontinuities. Hence, a dataset was composed with 488 marketing authorization application review procedures and reputational indicators of the applicants. All procedures received approval by the EMA between 1995 and 2017 and had comparable levels of product uncertainty. Using negative binomial and cox proportional hazard regressions allowed for validation that organizational reputation significantly affects the time to decision for products with comparable levels of uncertainty. Variables for all reputational dimensions showed statistical significant positive relationships with the duration of review times. In particular the number of current applications and proportion of failed applications of an applicant and general safety issues extended review times. Social judgements by the regulator appear to be mainly based on generalized favorability, as this was most evidently present in the data. For practice, this means that the EMA should be aware that even when relying on scientific information as base of decision-making, social judgements cannot be neglected. Additionally, future applicants can use these insights to submit new applications via subsidiaries, or in collaboration with others with a more favorable reputation.

Keywords: social; judgements; organizational; reputation; decision-making; EMA; CHMP; drug review

Acknowledgements

I would like to express my sincere gratitude towards my supervisor Jarno Hoekman, for his quick responses outside working hours, genuine feedback, our high-quality discussions about complex matters and his willingness to provide me with all sorts of data. Additionally, I thank Dr. Jacoline Bouvy of the National Institute for Health and Care Excellence (NICE) for providing me with data on product referrals and Ruben van Duijnhoven of the University of Amsterdam for making his research data on clinical trials sizes publicly available. I also like to thank my fellow students and teachers of the Innovation Sciences master for challenging and motivating me during the last two years. Lastly, I thank everyone who showed interest in my progress during the process of writing the thesis, especially my parents, whom provided me with the essential mental support during this stressful period.

Table of contents

List of abbreviations

AA	Accelerated Assessment
ATC	Anatomical Therapeutic Chemical
CIRS	Centre for Innovation in Regulatory Science
CHMP	Committee for Medicinal Products for Human Use
EMA	European Medicines Agency
EPAR	European Public Assessment Report
FDA	Food and Drug Adminstration
INN	International Nonproprietary Name
LoQ	List of Questions
MAH	Marketing Authorization Holder
MAA	Marketing Authorization Application
NAS	New Active Substance
SPC	Summary of Product Characteristics

1. Introduction

Before new pharmaceuticals are authorized to enter the market, their efficacy and safety have to be demonstrated by the organization that applies for market authorization. These 'applicants' submit an application at the regulatory agency that controls market entry in a restricted geographical area, such as the European Medicines Agency (EMA) for the European Union and US Food and Drug Administration (FDA) for the United States. The EMA and FDA demand extensive collection of pre-clinical and clinical data to assess product efficacy and safety, which leads to long development time and high costs of drug development (Jayasundara et al., 2019; Kaitin, 2010; Morgan et al., 2011).

Despite these prerequisites, regulators face inherent uncertainties when assessing these new drug applications for market entry, because not all data can be obtained in a timely manner (Carpenter, 2004). The uncertainty in assessing drug applications leaves the regulators with a dilemma between granting early market access and ensuring public safety (Eichler et al., 2008). This dilemma forms a 'stopping problem', which is defined as deciding at what moment to stop the review process and approve the new drug despite there still being uncertainties about the exact quality of the product (e.g. future side-effects) (Carpenter, 2002; Carpenter, 2004). Approving a new drug too quickly could result in reputational damage for the regulator when adverse side-effects occur after market authorization, while delaying the decision allows for additional safety assessments, but leaves patients untreated (Maor, 2011).

A variety of factors determine product uncertainty at the time of assessment, such as its pharmacological complexity, the amount of clinical data which is collected and prior use in other geographical areas (Moffit, 2010; van Duijnhoven et al., 2013). Moreover, the effect of product uncertainty on the speed of reviewing marketing authorization applications (MAA) has received considerable attention in scientific literature (Carpenter et al., 2010; Olson, 1997; Stern 2017). These scholars showed that not only the level of product uncertainty indicates the required review time, but social effects also play a part, leading to large differences in approval times between products (Alqathani et al., 2015; Roberts et al., 2011; Beishon, 2014; Downing et al., 2012; Downing et al., 2017).

To reduce uncertainty in the decision-making process and handle the stopping problem, regulators look at characteristics of the applicant, additionally to product features (Connelly et al., 2011; Noll, 1976; Spence, 1973). These 'social judgements' proxy the probable and unobserved quality of the product under assessment and therefore lower the uncertainty that the regulator faces (Tafuri et al., 2014). Prior studies emphasized the importance of three key characteristics to make social judgements of organizations, namely legitimacy, reputation and status (Bitektine, 2011; Jensen & Roy, 2008; Rao, 1994). While legitimation judgements are considered given in the context of drug approvals, since illegitimate organizations would not receive approval whatsoever, and the role of status has been studied before (e.g. Kim, 2012), the exact role of reputation in drug assessment and its effects on review time remain largely unclear.

Following Lange et al. (2011), I argue that regulators are affected by organizational reputation of applicants in three dimensions. Regulators can judge the reputation of applicants because the applicants are (1) known, (2) known for something or because the regulator has some impressions about the (3) generalized favorability of that company. Whilst being known emphasizes on familiarity between regulator and applicant, the second dimension uses previous performances as a likelihood estimation for future outcomes. The third dimension is a judgement that emphasizes on the overall attractiveness of the applicant (Lange et al., 2011). These dimensions have separately been studied in a variety of contexts (e.g. Rao, 1998; Rindova et al., 2005; Petkova et al., 2014). Some studies have specifically focused on the role of organizational reputation in the performance of firms in regulated markets, such as the banking industry (e.g. Deephouse & Carter, 2005; Englert et al., 2018).

Empirical studies on the highly regulated pharmaceutical industry also incorporated the effects of organizational reputation on drug review (Carpenter, 2002; Moffitt, 2010; Olsson, 1997). The specific effects of organizational reputation on drug assessment times have been investigated by several scholars (Carpenter et al., 2010; Kim, 2012; Olson, 1997). Their efforts focused unanimously on the FDA and oversimplified reputation, emphasizing that reputation is mainly acquired from either experience by previous interactions or generalized favorability. Consequently, currently no scientific work exists that investigates on how organizational reputational influences drug review times at the EMA. Even though Hauray (2017) and Tafuri et al. (2014) indicated that besides 'objective' knowledge through clinical trial data, social factors affect decisions by the Committee for Medicinal Products for Human Use (CHMP) as well, which is considered the decisive body of the EMA. This indicates social judgements about the reputation of applicants are used by the CHMP to reduce uncertainty about product quality.

By including the degree of product uncertainty, I follow Carpenter (2002) who argued that regulators face inherent uncertainty in decision-making due to unobservable product quality. By unravelling the concept of reputation, I test for effects that familiarity with the regulator, firm-specific prior performances and general safety events have on review times. Especially investigating organizational reputation based on prior performances in regulatory procedures has not happened in earlier work by Kim (2012), Olson (1997) or Carpenter et al. (2010). Because, contrary to the EMA, this data is mostly not publicly available for the FDA.

By applying the social judgement perspective by Bitektine (2011) to the context of new drug assessment, using the three dimensions of organizational reputation by Lange et al. (2011), I extend the body of knowledge that exists on the effects of social judgements in regulated markets. I do so, by empirically investigating the case of the pharmaceutical industry in Europe, focusing on the influence of the judgement of organizational reputation on review times of MAAs by the CHMP. Hence, the following research question applies: *To what extent do three different types of organizational reputation influence the length of the active review process of marketing authorization applications by the CHMP?*

Including the three dimensions of reputation provides new insights in the role that organizational reputation plays in decision-making under conditions of inherent uncertainty. Additionally, by focusing on the currently understudied EMA, allows to validate the notions made by Hauray (2017) and Tafuri et al. (2014) about the decision-making process of MAAs in Europe. Thus, I shed light on how expectations of the CHMP towards the applicant affect the speed of reviewing the MAA, as already happens for obtaining conditional market approval, where the CHMP considers the ability of the applicant to deliver future obligations (EMA, 2019).

From a societal perspective, this research is beneficial for pharmaceutical companies, patients and regulators of pharmaceutical markets. It highlights how the reputation of firms affects the length of the decision-making process. Shorter review times can be favorable for the applicants, because they possibly face comprehensive first-mover-advantages and have more time to exploit their patents and make return on investments (Stern, 2017). It can also be favorable for patients, who obtain quicker access to new medicines (Grabowski, 2002). Additionally, the EMA could become aware of unintentional differences in reviewing MAAs based on the reputation of applicants. This is especially relevant since the EMA highly values objectivity in product assessments. With this work, the EMA can decide if the effects of social judgements is detrimental and should be avoided, or whether it provides a useful heuristic in drug review. The following section explains the theoretical background for this research, followed by the methods to perform the study.

2. Theoretical background

2.1 Drug approval as a stopping problem

Unlike the producers, consumers of pharmaceuticals lack information to individually assess novel drugs on efficacy and safety. If not tackled by safety regulation, this would result in market failure because consumers are not adequately protected from product hazards (Green & Moore Jr., 1972). This is the rationale for having regulatory organizations that oversee market entry of new drugs (Kim, 2012). These organizations act as judges to ensure quality and additional social benefit of new drugs (Ehman et al., 2013). Regulatory agencies do so by installing numerous regulations that allow them to scientifically assess the measurable and observable characteristics of new drugs (Putzeist, 2013). However, just as individual consumers, regulatory agencies face inherent uncertainty and are limited by bounded rationality in decision-making, which implies they cannot make a perfect decision (Carpenter, 2004; Cooper & Kovacic, 2012). Hence, they cannot foresee possible negative outcomes of their decision, such as long-term adverse side effects or efficacy in combination with other drugs due to a lack of information (Mullins et al., 2010).

Ideally, regulators would wait with granting market approval until all data has been gathered to perfectly assess drug safety and efficacy, but they cannot because of the tender balance between gathering sufficient data and allowing quick market access to ensure medication for patients in need (Baird et al., 2014; Eichler et al., 2008). This dilemma is why regulatory agencies in the pharmaceutical industry face a 'stopping problem' (Carpenter, 2002; Carpenter, 2004). They must consider whether they have sufficient data to decide if the apparent benefits of the new drug outweigh the potential risks. Since the regulator is only bounded rational, there are two types of errors that might occur (Carpenter, 2004). A type I error is committed when a new drug is falsely considered safe and granted marketing authorization, while a type II error falsely rejects a MAA that should actually have been approved. The regulator does not want to wait too long with making a decision, nor does it want to approve an unsafe drug, because both harm its reputation (Carpenter, 2002; Maor, 2011). Hence, it is interesting to investigate the time a regulator needs for reviewing the drug, since this reflects at what point in time the regulator estimated that the remaining level of product uncertainty is acceptable for avoiding a type I error.

2.2 Product uncertainty

There are multiple sources of uncertainty of the product under assessment, mainly determined by what is already known about the (type of) product and its working mechanisms. Previous literature used a variety of indicators to capture product uncertainty, such as the targeted therapeutic category (Olson, 1997), priority designation (Moffitt, 2010) and entry order (Stern, 2017). However, most researchers capture the level of uncertainty by variables that distinguish between groups and thus are limitary distinctive, such as molecular size (e.g. Hofer et al., 2018; Regnstrom et al., 2010; Trusheim et al., 2010) or therapeutic class (e.g. Olson, 1997; Olson, 2004; Polidoro & Theeke, 2012). I argue that the level of product uncertainty can be approximated by three factors that express the amount of information that is currently present about the product and its features. These are the product complexity, evidence base and approval history.

2.2.1 Complexity

Product complexity is often included in studies on review time of pharmaceuticals and proved to significantly affect the length of the review procedure (Moffitt, 2010; Olson, 1997). Still, there is no agreement of what product complexity exactly entails, as scholars used divergent ways to approach product complexity and often confuse it with the novelty level of the product, by indicating whether the product consists of a known or new substance (Hofer et al., 2018; Stern, 2017). However, new substances can be very similar to existing products in terms of working mechanism, which indicates that information about the working mechanism is already available. But still, this does not provide

information about how complex the working mechanism actually is and thus using the product novelty level as a measure of product uncertainty seems inaccurate.

When looking at complexity of products, generally the distinction can be made between the complexity of the product structure and the complexity of its working mechanism. Product structure refers to the complexity of developing and producing the active substance and is often measured by distinguishing between small molecules, biologics and more complex substances like cell and gene therapies (Regnström et al., 2010; Tusheim et al., 2010). The former are often defined as chemically synthesized substances, while biologics are larger (protein) structures manufactured by means of biological sources (Tusheim et al., 2010). In general, small molecules are easier to produce and administer to humans than biologics.

Complexity of the working mechanism is based on the idea that some specific therapeutic categories are more complicated to treat than others (Olson, 1997). Often the therapeutic indication and entry order for a specific indication are used to indicate uncertainty levels (e.g. Kim, 2012; Stern, 2017). However, the number of therapeutic categories ranges from seven (Olson, 1997) to about twenty (Polidoro & Theeke, 2012) and entry order is often limited to first-, or later-in-class (Eder et al., 2014), which allows for limited distinction in complexity. Moffitt (2010) reasoned that product complexity is best captured by the length of the pharmacological description in the product label, since more complex products request a more extensive description of product features. Therefore, she argues that product complexity is best captured by looking at the length of the pharmacological description. This includes the amount of uncertainty regarding both its chemical structure and pharmacological properties and most accurately reflects complexity of both the structure and working mechanism (Moffitt, 2010).

2.2.2 Evidence base

Next to product complexity, the amount of knowledge about the quality of the product under assessment determines product uncertainty (Carpenter et al., 2010; Olson, 1997). When more information on product safety and efficacy is available, the regulator can make a more accurate estimation of the risks associated with using the product, which leads to lower chances of making a type I error. Previous literature used various measurements to estimate how regulators interpret the risks associated with a new product, based on the evidence flanking the product application (e.g. Carpenter et al., 2010; Kim, 2012; Kyle, 2006; Olson, 1997). For example, Olson (1997) uses the specialization of the applicant in making and selling pharmaceutical products as a proxy indicator to show how quality of the MAA influences review time at the FDA. However, this measurement is an indirect measurement and less likely to reflect the quality of the product under assessment and more the ability of the firm to deliver a high-quality and complete application (Olson, 1997).

Other scholars also used indirect measurements of the quality of the scientific evidence closely related to the MAA, such as citations analyses of scientific literature or patents (Carpenter et al., 2010; Kim, 2012; Kyle, 2006). Literature citations analyses rely on the assumption that products of higher quality are more likely to be mentioned in other scientific articles on product safety and efficacy (Kyle, 2006). Patent citation analyses of the underlying patents assume that reliable products are more likely to be built upon and therefore receive more patent citations (Kim, 2012). However, these indirect measurements have proven downsides such as that more literature citations might be related to higher safety concerns and the possibility of producing firms to generate favorable scientific publications (Kyle, 2006; Polidoro & Theeke, 2012). Patent citations have the downside that often the connection between highly cited patents and valuable inventions is lacking (Abrams & Sampat, 2017).

To include a direct measurement of the amount of evidence available, Kyle (2006) mentioned that frequency and severity of adverse interactions and side-effects also reflect the reliability of a drug.

Although products of higher quality are ought to have fewer side-effects, the occurrence of side-effects is during the application procedure inherent to the size of the clinical program, since larger trials will automatically result in more known side-effects (Eicher et al., 2008). Consequently, having more side-effects identified during the approval procedure actually lowers uncertainty about the drug, because the side-effects identified are included in the consideration to grant or deny marketing authorization. Thus, the size of the clinical program included in the application is more likely to directly reflect the amount of remaining uncertainty about drug reliability than the outcome of the program itself.

2.2.3 Approval history

The approval history of a product indicates product uncertainty, because when already approved elsewhere, a risk-benefit consideration is already performed by the regulators in that area (Carpenter, 2004). Since agencies similar to the EMA request likewise risk-benefit and efficacy and safety assessments, the EMA's opinion rarely deviates (Downing et al., 2012). For example, the FDA an EMA agree in their decisions on market entries in up to 98% of the cases (Kashoki et al., 2019). Additionally, if a product has already been granted market approval in another country, uncertainty about efficacy and safety is lower, since the product is already in use in the targeted population. On the contrary, if a product has no approval history at all, fewer clinical information is present and the EMA could be extra stringent is assessing the data presented.

2.2.4 Sources of product uncertainty

In sum, there are multiple sources of product uncertainty that regulators face in their decision-making, but uncertainty levels largely differ per product and MAA. Three sources that determine product uncertainty are considered here. These are the level of complexity, the evidence base and approval history of the product.

2.3 Firm signaling

Scholars have mentioned that regulators cope with uncertainty about product efficacy and safety, by relying on other signals that might give additional information about the product (Hauray, 2017; Tafuri et al., 2014). By assessing these alternative signals, the regulator tries to value the product and lower the 'stopping problem' they face. For example, regulators might take into account firm characteristics such as reputation (Olson, 1997) and status (Kim, 2012) which signal the sincerity of the applicant to the regulator. The signaling theory, originally developed by Spence (1973), proposes a useful starting point for understanding behavior of two parties facing information asymmetry, which indicates that the applicant has more information about product quality than the regulator. The general idea of the signaling theory is that organizations or individuals sent out signals to communicate their (unobserved) product quality to receivers (Basdeo et al., 2006; Bergh et al., 2014; Connelly et al., 2011). To comprehend how regulators perceive these signals, I firstly address how firms communicate these signals to regulators. Subsequently, I address how regulators make organizational judgements based on these signals and formulate hypotheses how this affects decision-making by the CHMP.

Most signaling literature has focused on actions that firms can perform to signal unobserved quality (Bergh et al., 2014; Connelly et al., 2011). Bergh et al. (2014) defines signals as '*observable actions that provide information about unobservable attributes and likely outcomes' (p.1335).* However, Olson (1997) showed empirically that it is not only the (timing of) actions a firm performs that sent signals to regulators, but that firm-specific characteristics also convey signals. She mentions: *'Firm-specific characteristics signal information to the reviewer about the expertise, experience, or reputation of the firm submitting the new drug application' (Olson, 1997, p. 385).* Hence, signals consist not solely of actions performed by the firm, but also its characteristics or appearance. It is to bear in mind that these firm-specific characteristics are derived from the past actions a firm has performed. For example, starting affiliations with high quality suppliers influences how others perceive the status of a firm (Benjamin & Podolny, 1999) and price setting determines reputation (Basdeo et al., 2006). Other

characteristics such as patent intensity, spending on research and development, or size also directly result of past actions and sent observable signals that affect decision-making, as Olson (1997) showed. To incorporate both organizational characteristics and actions the following definition is used to describe signals: *"observable firm characteristics or actions that provide information about unobservable attributes and probable outcomes".*

There are many ways in which a firm sends signals to a receiver, sometimes purposely, other times unwittingly. In the approval process with drug regulators, all signals, purposely or not, can be categorized by the features signal *immediacy* and signal *mode.* The *immediacy* of the signal refers to whether signals are conveyed directly from applicant to regulator, or through alternative ways. Direct signals are for example sent by applicants that apply for expedited approval pathways for orphan drugs, or have previously shown to develop drugs that meet unmet medical needs (Moffitt, 2010). Indirect signals communicate the knowledge and skills of the applicant to the regulator, through actions are not addressed for the regulator directly, but will be noticed, such as establishing collaborations (Stuart et al., 1999) or activating media and patient groups to show product necessity (Carpenter, 2002).

Signal *mode* refers to whether the signals received by the regulator, are sent through the product or the applicant during the ongoing MAA procedure. Product signals yield knowledge about the product undergoing the assessment procedure and determine the amount of uncertainty about the safety and efficacy of that particular product. Direct product signals are for example the size of the clinical program or applications at other regulatory agencies. In short, these reflect what information is reviewed by the CHMP to determine the safety and efficacy and constitute the level of product uncertainty. The direct product signals can be captured by the concepts complexity, evidence base and approval history as described in section 2.2. Indirect product signals are actions that the applicant can use to shape the acceptability boundaries regarding product approval. For example producing scientific publications independent from the clinical program (Polidoro & Theeke, 2012), exerting pressure on the regulatory agency by gaining media attention for the disease, mobilizing patient advocacy groups (Carpenter, 2004) or being the first to treat a particular disease (Carpenter, 2002). These actions influence the level of product uncertainty that is seen as acceptable to grant marketing authorization, at the time of decision-making.

Table 1: Typology of signals

	Product	Applicant
Direct	Signals that determine product uncertainty: • Is product under assessment already approved somewhere • Is a complete approval application with extensive clinical program submitted for product under assessment • Is product under assessment a novel treatment (for an untreated disease or according to new working mechanism) and is applied for expedited approval pathways	Signals that provide information about the quality of the applicant based on own experience: • Contact with the regulator through running product applications • Previous product approvals • Previous safety scandals • Previous ability to serve customers
Indirect	Signals that shape acceptability boundaries of product uncertainty: • Activating media and patient groups to generate attention for drug currently in approval procedure • Produce scientific publications to raise awareness among professionals about drug currently in application procedure	Signals that provide information about the quality of the applicant based on the experience of others: • Previous and current affiliations with high-status organizations within the industry • Previous and current specialization in disease area and product type

Applicant signals refer to actions unrelated to product being reviewed in the running MAA. Direct applicant signals communicate predicted performance based on previous interactions between the regulator and applicant, showing that the organization is experienced and whether it previously has been able to produce quality products (Deephouse & Carter, 2005). Indirect applicant signals communicate organizational knowledge that might predict to what extent the sender's products are of high quality, based on interactions with others in the industry, but not directly aimed at the signal-receiver. For example mergers or other long-term collaborations indicate ameliorated in-house knowledge about specific disease areas or product types (Podolny, 1993). These ways of sending signals have been summarized in table 1. Now it has been clarified how direct and indirect product and applicant signals influence (the acceptability of) product uncertainty, I argue how a regulator perceives these signals, based on the theory of social judgement of organizations by Bitektine (2011).

2.4 Social judgements of organizations

In open markets firms sent signals directly to their customers to influence their judgements about the organization, but in regulated markets the regulator makes this judgement for them (Bitektine, 2011). Regulators are ought to make more rational judgements than consumers because they consider the high stakes regarding product risks and benefits, objectively from personal need. To investigate how regulators make these 'rational' judgements, I use the ideas on social judgements of organizations by Bitektine (2011). He argues that regulators under conditions of uncertainty, as is the case in the pharmaceutical market, receive signals about the legitimacy, reputation and status of the organization. I adopt an organizational perspective where legitimacy, reputation and status judgements of organizations are not seen as competing processes where one form of judgement is dominant, but rather as complementary features of the firm under assessment, as is often proposed in scientific literature (Cattaneo et al., 2016; King & Whetten, 2008; Rao, 1994; Washington & Zajac, 2005). This means that the legitimacy, reputation and status of an organization is always judged, but the objective of the evaluator determines what the decisive judgement is in the decision-making. The remainder of section 2.4 is focused on clarifying these concepts as they are often confused and merged in scientific literature (Deephouse & Suchman, 2008; Piazza & Castellucci, 2014).

2.4.1 Legitimacy

The definition of legitimacy depends on both the scientific background that is chosen and who is regarded as the evaluating audience (Deephouse, 1996; Bitektine, 2011). Here, organizational legitimacy is defined according to Deephouse et al. (2017), who state: "*Organizational legitimacy is the perceived appropriateness of an organization to a social system in terms of rules, values, norms, and definitions" (p.7)*, where the social system in the pharmaceutical industry consists of the regulator and the general public. The concept of organizational legitimacy is often divided into several (closely related) types (Ruef & Scott, 1998; Suchman, 1995). The most common distinction has been made by Aldrich and Fiol (1994), who differentiate between cognitive and sociopolitical legitimacy.

Cognitive legitimacy, according to Aldrich and Fiol (1994), is a category-based judgement that refers to whether the existence and activities of the firm are 'taken-for-granted'. The cognitive legitimacy of a firm or product is based on whether the firm is conforming to a predefined category or previously encountered form (Deephouse, 1996; Deephouse & Carter, 2005). Sociopolitical legitimacy refers to whether the organization is seen as legitimate, compared to the prevailing social norms (Aldrich & Fiol, 1994). However, the legitimation judgements made by regulators in regulated industries are dichotomous decisions that either allow or reject firms to enter the market, and thus determine survival of organizations (Bitektine, 2011). The outcome of the legitimation judgement is that the behavior and form of the organization are either in line, or not in line with known cognitive categories (cognitive legitimacy) and social norms (sociopolitical legitimacy).

For example, a company with an organizational form that has never been encountered, has no cognitive legitimacy and would therefore not be allowed to the market (Bitektine, 2011). In contexts adjacent to the pharmaceutical industry, empirical studies have operationalized cognitive legitimacy by measuring accreditations, showing conformance with regulatory norms (Ruef & Scott, 1998) or conformity with regulatory standards (Deephouse & Carter, 2005). Sociopolitical legitimacy of firms in a pharmaceutical context is mostly reflected by whether the products they produce overcome unmet medical need. If the targeted population is not present in Europe, or the population is small and cost-benefit ratio highly transcends acceptable norms the producing organization is regarded as illegitimate and rejected market access (Tafuri et al., 2016). Therefore, I argue that all organizations that have received marketing authorization by the CHMP have received positive sociopolitical and cognitive legitimacy judgements and consequently, having legitimacy is prerequisite for pharmaceutical companies with marketed products.

2.4.2. Reputation

There is some commonality between the concepts of legitimacy and the reputation of a firm since both are derived from the prevalence and the history of the particular form of the organization (Bitektine, 2011). Deephouse and Carter (2005) emphasized on the differences between these concepts: where legitimacy is an institutional perspective focused on comparing the organization to social norms, reputation is an economic perspective that highlights the expected quality of the producing firm. Reputation of the firm is established based on a relative comparison with other firms in the same industry and is especially prominent in markets where product quality is uncertain before purchase (Allen, 1984). This makes reputation especially relevant in the pharmaceutical industry, since regulators cannot completely assess product quality before large scale usage and are therefore likely to be reliant on previous performances by the organization.

Lange et al. (2011) showed in their literature review that reputation is a three-dimensional conduct, consisting of familiarity with the organization, perceived predictability and perceptions about the favorability of the organization. They have called these dimensions (1) being known, (2) being known for something and (3) generalized favorability. When an organization is well known, signal receivers have a strong, enduring sense of the distinctiveness of firm attributes, relatively to other firms (Whetten & Mackey, 2002). Being known for something refers to the subjective assessment of a particular attribute or characteristic that allows for a judgement regarding likely future performances (Lange et al., 2011). Generalized favorability is described as the cultural system that the observers and observed are immersed in, which socially constructs the standards for judging corporate favorability (Rao, 1994). It is an overall assessment that transcends specific aspects of the past of the organization.

These dimensions are distinctive in how they address the judgement of the signal receiver and what the judgement is based upon (Lange et al., 2011). Being known is a nonevaluative impression held by the receiver, about the whole of the organization. Being known for something is a judgement that relates to the organization or its perceived outcomes based on particular organizational attributes. This is the most specific judgement about the organizational reputation that is constituted, as it is both judgmental and particular for that organization. Lastly, generalized favorability is a judgement based on the generalized perception of the firm as a whole and not necessarily specific for one particular firm. Although these concepts appear highly similar, they are not necessarily correlated. For example, being known and generalized favorability are distinct in their level of judgement. An organization with little familiarity (being known) can still be liked or disliked within the group of signal receivers. Being known for something and generalized favorability are distinctive in the way that the former judges (un)desired specific organizational attributes to determine expectations, while the latter reflects generalized, global perceptions that determine the reactions to approach, or avoid interactions. Being known and being known for something are different in both level of judgement and scope. To clarify these dimensions, table 2 sets out how they are distinct on their type of judgement and focus.

Table 2: Three dimensions of organizational reputation, from Lange et al. (2011) p. 166

	Being known	Being known for something	Generalized favorability
Generalized or particular (focus on specific organizational attributes or characteristics versus generalized impression of the organizational whole)	Generalized	Particular	Generalized
Nonevaluative or judgmental (judgement by signal receiver about organization or outcomes versus impressions without judgement)	Nonevaluative	Judgmental	Judgmental

2.4.3 Status

There has been discussion about the definition of status, also called prominence (Rindova et al., 2005) or prestige (Jensen & Roy, 2008) and whether it is a signal of product quality, just as reputation (Jensen et al., 2011; Podolny, 1993). But although status and reputation are partly build up from the same antecedents, they differ specifically in their conceptualization. Both status and reputation allow judgements regarding the expected outcome, but where reputation is a judgement based on past performances, status is a judgement about how well the organization is, based on the perception of related organizations. Bitektine (2011) mentioned that social judgements are always a consideration between making your own based on earlier experiences, or adopting someone else's and that status judgements are especially based on perceptions of others. In other words, status judgements are not based on the known past performance, but by its current quality potential derived from the pattern of affiliations and how relationships with other actors are maintained (Rindova et al., 2006). The underlying assumption for this reasoning is that if others are willing to affiliate themselves with the firm, the quality of the firm should be good (Podolny, 1993).

Empirically, status has been operationalized in a variety of ways, mainly focusing on the relationships that an organization hold towards others. For example, mergers and acquisitions (Ruef & Scott, 1998), affiliations formed with other organizations of high status (Stuart et al., 1999) and affiliations with high-status educational institutions such as highly certified employees (Philips & Zuckerman, 2001) have been used to indicate status. Other used measurements were media rankings, certification achievements and ties to high-status actors (Rindova et al., 2005) or centrality, based on co-investment connections (Petkova et al., 2014). These ways of operationalizing status all reflect how willingly other parties in the industry are to engage with the organization that is receiving the status judgement. In a pharmaceutical context, status measurement focused on visualizing ties between organizations in several ways. Stern et al. (2014) highlighted status and reputation in alliance formation between pharmaceutical companies and biotech firms by looking at personal scientific output (reputation) and attended universities using academic rankings (status). But since these measurements are on individual level, they are inappropriate when considering impact of firm characteristics. Hu et al. (2015) and Kim (2012) focused on status by looking into patent ownership and citations, which can be used to show connections between different actors in an industry.

2.4.4 Demarcating reputation and status

Although reputation and status are used as judgement to decrease uncertainty of product quality ex ante and are both derived from the historical actions an organization has performed, they also differ in several dimensions (Sorensen, 2014). The most important distinction is that reputation is granted by evaluating direct earlier contact with the firm that holds the reputation. Through consumption and interaction one can know whether the products or services delivered by that firm are in fact of high quality. Status on the other hand is a social construct that values individuals or organizations based on their position in a social hierarchy (Patterson et al., 2014). This means status judgements do not involve valuation of direct contact with the one that holds the status. To highlight the differences between reputation and status, table 3 gives an overview of these concepts and summarizes how operationalization in contexts adjacent to the pharmaceutical industry has been.

Table 3: Differences between organizational reputation and status

Dimension	Reputation	Status
Scientific origin	Economic literature	Organizational and research management literature
Cognitive reasoning	The perceived ability of the firm to produce quality products based on own previous consumption of products or interaction with the firm.	The position of the firm in the social hierarchy defines to what extent others trust the quality of future products by that firm.
Firm evaluation by	The reputation of the firm is compared to the reputation of all other firms in the industry, often based on earlier experiences.	The status judgement relies on how the social position of the firm is appreciated, without using own prior experience or interactions as measure of assessment.
Relevant empirical measurements in adjacent contexts	• Certification contests (Rao, 1994) • Content analysis of media publications (Deephouse, 2000; Englert et al., 2008) • Expert ratings (Rindova et al., 2005) • Prior product approvals (Kim, 2012) • Number of ongoing MAAs (Olson, 1997; Stern, 2017)	• Alliances or affiliations with actors of high status (Stuart et al., 1999) • Affiliations with high-status institutions (Philips & Zuckerman, 2001) • Co-authored scientific publications (Hu et al., 2015) • Patent citations (Kim, 2012) • Number of owned patents (Hu et al., 2015)

2.5 Judgements of legitimacy, reputation and status in pharma

When taking the theory of social judgements of organizations to the context of the European pharmaceutical industry, the CHMP is the executive power of the EMA and the signal-receiver who makes judgements about organizations that have submitted MAAs. As already mentioned in section 2.3, the judgement of legitimacy, reputation and status are competing processes, but are not all equally relevant when looking at the effect of organizational judgements on review time. In the pharmaceutical industry, a positive legitimation judgement is prerequisite for market access and fundamental to the survival of an organization. Hence, the judgements about reputation and status are of more interest when looking at the review process of new drugs. As already mentioned, status judgements are not derived from the CHMPs own experiences with the applicant, but are based on the perception that the CHMP has about how others in the industry value its knowledge or qualities.

I argue that the judgement of the reputation of the applicant is most decisive in drug review and especially when looking at the European context. Jensen and Roy (2008) mentioned that both reputation and status judgements are used in the decision-process as a two-staged process. Where status judgements are the first stage and allow to screen for all possible alternative exchange partners that are to be considered, the reputation judgement is involved in the second-stage and affects choosing amongst the set of alternatives acquired after the status judgement (Jensen & Roy, 2008). Bitektine (2011) adopts the ideas of Jensen and Roy (2008) and emphasized that status, which is an efficient category-based judgement, happens prior to the more complex judgement about reputation. However, as the EMA is obligated to review MAAs submitted by any European firm that is able to meet the financial requirements, it does not use status judgements to decide which applicants it reviews. This is because it has no possibility to choose between a set of alternative exchange partners, but is obligated to review products by any applicant, independent from its status. Therefore, the status judgement likely affects decision-making by the CHMP to a lesser extent than the reputation judgement.

Nonetheless, while reputational judgements are probably the decisive factor for the CHMP, status judgements do still play a role and thus cannot be neglected. Kim (2012) showed that status of applicants is particularly relevant because it works as a buffer for threats to the reputation of the regulator. This means that regulators could be willing to grant faster approval for applicants with a higher status, because the status of the applicant defends the regulator for potential reputational damage after a type I error.

2.6 Reputation in reviewing MAAs

There is a considerable body of knowledge that has focused on the effects of organizational features on regulatory decision-making in a pharmaceutical context, specifically considering the FDA. Many have included some measurement of reputation, but have not distinguished between being known, being known for something and generalized favorability, this section is dedicated to classifying their efforts according to the dimension of reputation as proposed by Lange et al. (2011), as summarized in table 4.

Table 4: Empirical contributions that included reputation measurements

Author(s)	Being known	Being known for something	Generalized favorability
Olson, 1997	Previous approvals Current applications	-	Research intensity Specialization
Carpenter, 2002	Previous approvals	-	Entry order Specialization
Carpenter et al., 2010	Previous approvals	-	Entry order Disease attention
Moffit, 2010*	Previous MAA submissions	-	Disease attention
Kim, 2012	Previous approvals	Product safety withdrawals by specific applicant	Entry order Industry safety events
Stern, 2017	-	-	Entry order

* Focus was not on FDA review times, but estimated chances of MAA approval under several circumstances

As touched upon in section 2.4.2, reputation consists of three dimensions. The first being whether the organizational is known among the reviewers. This dimension deals with the familiarity between applicant and regulator, or the extent to which the regulator is aware of the applicant. The first scholar to focus on predictors of FDA review times was Olson (1997; 2004), who measured reputation by the experience the applicant had gathered over time, focusing on whether an organization was known from previous interactions. She counted the number of past approvals and ongoing applications as measurement of experience with the regulator and found that especially the number of ongoing MAAs showed to strongly reduce review times at the FDA. Scholars of similar contributions tested the same relationships and included past or current interactions with the regulator as measurement of the experience of the applicant (Carpenter, 2002; Carpenter et al., 2010; Moffit, 2010; Kim, 2012). However, only Carpenter (2002) validated her results by finding significant negative effects of the number of previous approvals on review times of MAAs. That other studies have not found these effects is theoretically expected, as being known constitutes a non-evaluative judgement according to Lange et al. (2011).

The second dimension of reputation, being known for something, has to a lesser extent been investigated in prior work. None of the scholars that focused on FDA review times have included the specific prior interactions of the applicant with the regulator that yield future expectations, except for Kim (2012). He included whether an applicant previously was involved in any safety withdrawal, but found no significant effects on review times. Hence, there is a lack of knowledge on the attributes that allow for an evaluative organization-specific judgement of the applicant. This is especially worth mentioning, because according to Lange et al. (2011) the judgements of what firms are known for are primarily used by the regulator to constitute future predictions.

The third dimension, the generalized favorability is included in several studies, but in distinct ways. Olson (1997; 2004) included both the research intensity and the level of specialization as characteristics that are likely to influence review times. She argued that expectations of regulators are more favorably when applicants show commitment to research or being specialists or experts. Her data showed that both indicators significantly decreased review times. Others included the entry order of the drug or the attention for a specific disease as the generalized favorability, hypothesizing that

MAAs that treat an underserved disease area receive faster approvals (Carpenter, 2002; Carpenter et al., 2010; Kim, 2012; Moffit, 2010; Stern, 2017). Lastly, Kim (2012) argued that safety withdrawals have repercussions on the generalized favorability of the whole industry. He found evidence that major adverse events affect the favorability towards all future applicants, even more so than they harm the reputation of only the producing company.

2.7 Determinants of review time

In the previous sections the three dimensions of reputation are discussed and I have clarified how prior scholars have researched the effects of these types of reputation on review times in the context of pharmaceuticals. This section is dedicated to hypothesizing how organizational reputation affects the time the CHMP needs to review a MAA.

2.7.1 Being known

According to Lange et al. (2011) being known is a nonevaluative reputational judgement and a general perception about the applicant that is derived from previous encounters. It is not distinctive in outcome whether the previous encounters had positive or negative outcomes, but reflects the familiarity amongst the regulator. As all scholars who focused on product review times at the FDA emphasized, both the number of previous approvals and current applications reflect the familiarity between regulator and applicant (Olson, 1997; Carpenter et al., 2010). Although the judgement about being known is theoretically objective and subsequently there should be no effect on review time, there are several arguments that explain why having past experience could lead to faster reviews. Nonetheless, these arguments are not focused on explaining why more reputation by being known does lead to faster, since this is theoretically not distinctive. They rather explain how non-market capabilities that the applicant develops when gaining experience are likely to influence the rate of reviewing MAAs.

Both Olson (1997) and Carpenter et al. (2010) have included the number of past approvals in their analyses about drug review times at the FDA. They argued that prior approvals reflect the familiarity of the applicant with the regulatory process and simultaneously makes the regulators acquainted with the applicant. The process of receiving market approval for new products is complex, consisting of a plurality of regulatory access pathways, pathway facilitators, continuous changes, additions of procedural requirements and extensive demands for clinical trials (Baird et al., 2014; Eichler et al., 2008; Elsäßer et al., 2014; Liberti et al., 2017). For applicants that are unfamiliar with this procedure, the complexity makes it more difficult to submit complete applications and obtain marketing authorization as quickly as possible. Having experience in completing the procedure indicates the applicant's knowledge about the reviewing process, but also proxies the degree of familiarity between regulator and applicant (Olson, 1997). Besides learning effects from knowing the procedure, frequent contact can be advantageous for the applicant because it could result in interpersonal relationships between regulator and applicant (Bonardi et al., 2006). Hence, applicants that have built up experience from past approvals are likely to submit more complete and coherent applications and take advantage of established interpersonal relationships, which makes them better in quickly receiving marketing authorization. This results in the first hypothesis: *Hypothesis 1: The more previous approvals the applicant has had, the quicker its MAAs are being reviewed given comparable levels of product uncertainty.*

There are multiple reasons to believe that besides being known for having past approvals, it is also important to consider current experience of the applicant in obtaining marketing authorization (Olson, 1997). Firstly, regulations and procedures can change over time and past experience might therefore become irrelevant if too many changes have occurred. In contrast, having ongoing applications allows for adaptation and quickly becoming familiar with new regulations. Second, having more ongoing interactions with the regulator could lead to more frequent use of information channels via

interpersonal relationships between the applicant and regulator, which makes that expectations are more clearly communicated (Olson, 1997). If the expectation of the CHMP are communicated well, the applicant has the chance to gather additional data on possible bottlenecks or temper expectations before submitting the answer to the list of questions the CHMP compiles. Lastly, having interactions with the regulator could lead to reputational advantages since the representatives of the applicant are already known among the reviewers. These three reasons are likely to positively affect the reputational judgement by the CHMP and speed up the review process, leading to the second hypothesis: *Hypothesis 2: The greater the current experience of the applicant, the quicker its new product applications are being reviewed given comparable levels of product uncertainty.*

2.7.2 Being known for something

When applicants are known for some specific trait, attribute or past performance, this affects the reputational judgement that the regulator makes about that particular applicant (Lange et al., 2011). Few scholars researching drug review, have emphasized on the individual past performances of applicants and its effects on future review times. Only Kim (2012) accounted for firm-specific performances, testing how previous safety issues affected future review times. He found no significant relationship between the number of safety withdrawals caused by the applicant and review times. Theoretically, this dimension is most distinctive in the type of reputational judgement, because it is both subjective and specific for one applicant (Lange et al., 2011). It is the result of evaluations by the regulator on the likelihood that the outcomes are in line with the specific needs of the regulator. The needs of the CHMP are to facilitate timely access to medicines and product safety. This indicates that there are mainly two areas that allow judgement of applicants regarding what they 'are known for'. Applicants have a specific performance regarding the safety or the products they have produced and a specific performance based on whether they have provided access to medicines.

Kim (2012) emphasized that quality issues, such as product withdrawals provide an external shock that harm both regulator and the producer of the defective product. After quality issues, regulators are extra wary of making another type 1 error and likely hold suspicion towards individual applicants that have been engaged in these issues (Carpenter, 2002; Maor, 2011). This is, because they do not want to harm their own reputation again. Because the regulator estimates the likelihood of a(n) (un)wanted outcome on the current MAA based on previous performances by its producing company, it is likely to be reflected in review times. When an applicant has had quality problems with its products, this decreases the estimated likelihood in the eyes of the CHMP that the outcome of the current MAA is positive. Subsequently, having been involved in on average more scandals decreases the estimated likelihood of a positive future outcome. Hence, the third hypothesis reads: *Hypothesis 3: The more quality issues applicants have previously been involved in, the slower its MAA is reviewed given comparable levels of product uncertainty.*

A second reputational judgement of the applicant can be made based on how often it is involved in product discontinuities. Primarily, there are two main reasons why organizations fail to continuously deliver products to the intended patient population. Firstly, their product is withdrawn because it lacks the quality to be a long-term solution for the intended patient population. Either better products have been developed and their product is not used anymore, or the product has been approved without a clear patient population that the product was intended for. The second reason is that the application that has been submitted by the organization fails, because it lacks the required clinical data, or the data show that the product is not of such quality that the CHMP is convinced to approve the product. Either one of these options is likely to negatively affect the reputational judgement of the applicant.

When a product is approved by the EMA, it receives market authorization for five years, after which its marketing authorization holder (MAH) can apply for renewal. When a product is not commercially viable anymore due to better alternatives, or because the patient population disappears (e.g. vaccines which target a specific type of influenza), MAHs generally do not apply for renewal. For products that

have been withdrawn before the marketing authorization needed to be extended, the regulator starts to doubt whether the applicant will know its intended patient populations in future product applications (Hauray, 2017). I reason that companies that withdraw their product within five years after receiving marketing authorization, came up with products that have no clear patient populations or do not outperform alternatives (excluding voluntarily withdrawals or non-renewals that have been linked to safety or quality issues). Hauray (2017) confirmed that the CHMP makes judgements about applicants based on the suggested possible patient population of the submitted product.

There is not much scientific knowledge about the influence of product refusals in the pharmaceutical context, since these are not made public by the FDA and most scientific contributions have focused on the FDA. Basically, there are two scenarios in which products that are submitted for receiving marketing authorization fail. The first scenario being when the clinical evidence regarding the efficacy or safety of a product is not profound enough in the eyes of the applicant or the products seems insufficiently commercially viable, the application is often withdrawn before receiving an opinion by the CHMP. Sometimes the applicant thinks the evidence is strong enough but the CHMP refuses marketing authorization and issues a negative opinion. Four reasons can be thought of when considering product refusals, either there is not sufficient information about the efficacy and safety of the product, a better alternative is readily available, the application itself has not presented the clinical evidence in the right way or there is no clarity about who will use the product.

Both withdrawn applications and refused products can negatively influence the reputation of the applicant, since it possibly makes the CHMP doubt whether the applicant prospectively is able to produce products and submit applications of the required quality. When an applicant is associated with many previously failed applications, the CHMP is likely to question the quality of the application that is currently being reviewed. These expectations could lead to more extensive product reviews, because the CHMP is likely to have more questions about the usage of the product and the intended patient population. Hence, applicants that are associated with a high rate of product withdrawals or failed applications are likely to receive negative reputational judgements. Summarized, this leads to the fourth hypothesis: *Hypothesis 4: The more often an applicant has been involved in product discontinuities, the slower its MAA is reviewed given comparable levels of product uncertainty.*

2.7.3 Generalized favorability

Generalized favorability refers to the general perception of the applicant, while 'being known for something' referred to particular expectations regarding the specific outcome. The favorability of applicants of MAAs can roughly be divided in environmental factors that affect the perceived reputation and characteristics that tell something about the applicant in general, without inducing expectations based on previous outcomes. Here, the expected effects of major quality issues and the degree of specialization of the applicant are hypothesized

Both Carpenter (2004) and Kim (2012) have made the argument that reputation of all companies in the industry is affected by major safety issues. Kim (2012) argued: "*Adverse events also have ramifications that extend beyond the specific defective product or producer" (p.136).* Kim (2012) empirically showed this by emphasizing that product withdrawals in the year prior to new approvals significantly slowed down review times at the FDA. The reasoning is that because regulators are extra wary of making another type 1 error after quality issues (Carpenter, 2002; Carpenter et al., 2010). Regulators install new regulations or additional requirements for presentation of clinical evidence, as happened at the FDA in the 1960s after the thalidomide tragedy or after Vioxx was taken off the market (Maor, 2011). These regulations extend review times because new requirements are not part of the review routines and additional data assessment is time consuming. Installing these new measures reflects that the regulator tries to minimize the chances on future quality issues. These effects are

especially profound soon after quality issues, because new measures are not part of the organizational routines yet and the regulator's fear for repetition is most strong.

Besides safety withdrawals as adverse events, product suspensions after a risk-benefit reassessment also indicate safety issues. When serious adverse side-effects occur or the efficacy of product is doubted, the EMA starts a risk-benefit reassessment (Bouvy et al., 2016). When a centrally authorized product is reassessed by the CHMP through an article 20 procedure, doubts exist about whether the risk benefit ratio of the product is deemed positive enough. When a product is suspended after a risk-benefit reassessment, this indicates serious uncertainty about the quality of the product. Product suspensions harm the reputation of the regulator, who shall behave similarly as ex post a safety withdrawal, operating with more cautiousness out of fear for harming its own reputation again (Carpenter, 2002). Hence, both product suspensions following reassessment and product safety withdrawals are expected to affect the generalized favorability of future applicants. This reasoning is captured by the following hypothesis: *Hypothesis 5: When general safety issues happen shortly before the start of reviewing the MAA, the time for reviewing the MAA shall be longer given comparable levels of product uncertainty.*

Lastly, Olson (1997; 2004) showed empirically that the degree of specialization significantly shortens review times at the FDA. She argued that applicants with a higher degree of specialization can be considered experts in the production and sales of pharmaceutical products. This follows the logic that the sales by pharmaceutical products compared to the total sales indicates if that applicant is expected to make more specialized products (Harford et al., 2009). Being expected to produce specialized products expresses the perception about the expertise of the organization in delivering high-quality products (Rhee & Haunschild, 2006). Additionally, Jensen and Roy (2008) argued that specialization is generally associated with higher technical skills and business integrity. Shortly, being seen as an expert in the production of pharmaceutical products is likely to elicit a more favorable attitude of the regulator towards the applicant, which leads to a more favorable reputation judgement. Following, the sixth hypothesis reads as follows: *Hypothesis 6: The more specialized the applicant is in producing pharmaceutical products, the faster its MAAs are reviewed given comparable levels of product uncertainty.*

To summarize this chapter, figure 1 visualizes how product features determine review time and the hypothesized effects of reputation.

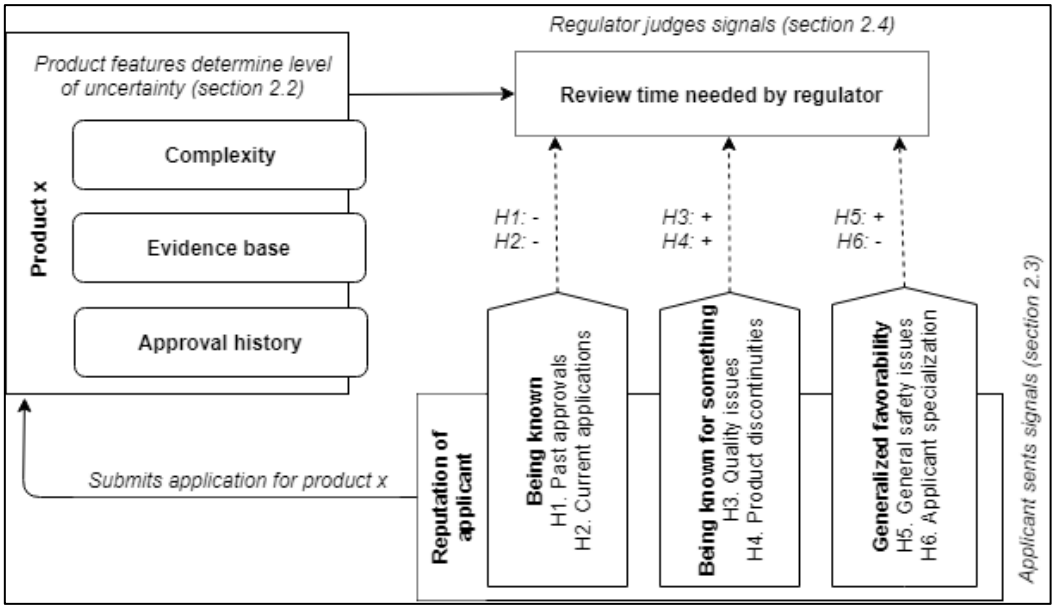


Figure 1: Visualization of the composition of review times based on product features and the different reputational signals which are expected to influence MAA review time

3. Methods

3.1 Research design

The deductive explanatory nature of this research requests a quantitative cross-sectional approach to answer the research question and achieve the research aims. Which is to give general insights into the effects of organizational reputation on the length of reviewing a MAA, after accounting for the uncertainty of the product under assessment. Second, by using the three dimensions of reputation as proposed by Lange et al. (2011), I add to the existing knowledge on regulatory decision-making in the pharmaceutical sector, by being the first to research the effects of three dimensions of organizational simultaneously, using multiple indicators on each dimension. Lastly, this research provides insights in the effects of organizational judgements on the length of decision-making at the EMA's decisive body, the CHMP, specifically focusing on the reputation judgement.

This research investigates the case of the European pharmaceutical industry by researching MAAs that received a positive opinion by the CHMP, since the establishment of the EMA in 1995. The case of the pharmaceutical industry is especially interesting since regulatory-decision making in drug approvals is subject to high levels of uncertainty and therefore regulators are more likely to rely on alternative sources of information. Olson (1997) and Kim (2012) already proved that firm features such as their reputation determine the length of review times at the FDA. Despite the FDA being of a different organizational setup, it equivalently regulates the marketing authorization of new drugs and performs an executing function similar to the EMA. Many studies have focused on the differences in drug approval times between the FDA and the EMA (e.g. Alqahtani et al., 2015; Makuch & Shi, 2014), but unlike the regulatory process in the US, hardly any quantitative studies have investigated the origin of varying drug review times within Europe. To shed light on the aforementioned, this research analyses the review process of new pharmaceuticals in the European market and focuses on the possible effects that social judgements of organizational reputation has.

3.2 Sampling strategy

Purposive sampling is used to gather the sample, consisting of all MAAs and their applicants that received market approval via the centralized market authorization procedure of the EMA. The timespan is limited from the founding of the EMA on the first of January 1995 until the 31st of December 2016. Initially, the set had been limited to only products that were approved through article 8.3 of the European Commission (EC), which requests full (pre-)clinical studies on pharmacodynamics, pharmacokinetics and toxicology, as obligated in module four and five of the Annex of Directive 2001/83 (European Commission, 2001). Limiting the set to only article 8.3 excluded all applications via article 10, consisting of generics, hybrid or similar biological products, which are subject to lower levels of uncertainty, because their working mechanism, efficacy and safety profile is already proven (EMA, 2019).

However, since the Directive 2001/83 of the EC only came into force in November 2001, this inclusion criteria does not cover all products from years before 2002. To include products with similar levels of uncertainty as approvals via article 8.3 applications, the European Public Assessment Reports (EPARs) were manually assessed and products were categorized as follows. Product applications that were submitted with full studies on pharmacodynamics, -kinetics and toxicology and conducted first-in-human trials in healthy volunteers were considered similar to approvals via article 8.3. Applications that substituted data from (pre-)clinical studies with existing literature were considered having lower levels of uncertainty and are therefore excluded. For applications that included all obligatory pre-clinical and clinical studies but did not mention first-in-human trials in healthy volunteers, a list of 'new active substances' (NAS) composed by the Centre for Innovation in Regulatory Science (CIRS), was used to determine uncertainty levels. If trials in healthy volunteers were not mentioned, but all other

necessary trials had been conducted, a product was only regarded as being of similar uncertainty as products via article 8.3, if CIRS classified the product as being a NAS.

During the next steps all vaccines (classified with Anatomical Therapeutic Chemical (ATC) code J07), radiopharmaceuticals (ATC code V09 and V10) and 'cloned applications' were removed. Vaccines and radiopharmaceuticals are subject to distinct regulations for proving efficacy and safety as stated in Directive 2001/83 (European Commission, 2001) and excluding these products is coherent with the definition of a NAS as established by CIRS, ensuring consistent sampling. 'Cloned applications' are considered all applications for the same compound submitted for the same therapeutic indication with different brand names. A cloned application is defined as: *Applications with either the same applicant and active substance (International Nonproprietary Name (INN)), ór different applicants with the same INN and same therapeutic indication, submitted maximum one year apart.* To remove all cloned applications, the list of applications was manually examined and all applications that met the definition of a cloned application were removed. In case of a cloned application, only the application with the earliest validation date was included. Lastly, seventeen products that received a negative opinion in the standard procedure, but were granted market authorization after re-examination were excluded from the sample set, because the review times of these products are largely dependent on the re-examination procedure and on average twice as long because the assessment procedure is repeated.

3.2.1 Applicants

For each application, the applicant of the MAA was extracted from the EPAR. The applicant is defined as: *The largest organizational entity that has been named by the EMA in the EPAR of the new product, as applying for marketing authorization of the new product, at the time of MAA submission.* Next, predictor variables have been measured for the applicant only, meaning that when drugs were submitted in collaboration with other organizational entities, reputation is not assumed to accumulate or being judged based on the average reputation of both applicants. When mergers or acquisitions between organizations occurred, reputation measured by applicant-specific variables was assumed to evaporate, which has the underlying assumption that the CHMP has no memory or reputational judgements about either one of the organizations involved.

3.3 Dependent variable

The sampling strategy described in section 3.2 resulted in a final sample of 488 approved products, as stepwise visualized in figure 2. The remainder of this section is focused on clarifying the dependent and independent variables, explaining how the independent, dependent and control variables have been operationalized and how data gathering for each variable has taken place.

Active review time. The product review time is measured by the number of days that the CHMP has actively reviewed the product, since this most accurately displays the time needed by the CHMP to review the drug application. The product assessment period of products applied via article 8.3, consists of *active review time* and *clock stop*, together making up the *total review time*, as displayed in figure 3. The total review time is considered the number of days from the submission of the application until the opinion of the CHMP. Since total review time is highly dependent on clock stop time, which is mostly reflecting the resources and efforts the applicant uses to reduce remaining product uncertainty, the active review time is a more accurate measurement when measuring effects on the length of the CHMPs decision making process (Beishon, 2014).

For earlier research, Hoekman & Boon (2019) already extracted review times and applicant companies from the documents with procedural steps and EMA annual reports. For these products the total review time, clock stop and active time were manually extracted from the EMA annual reports and checked with the data provided by Hoekman, to ensure no mistakes in the data. Subsequently, clock stop and active review time were compared with the total review time as given in the annual reports by the EMA. If any inconsistencies occurred when comparing clock stop and active time with the total

review time, dates were recomputed manually by extracting all dates from the procedural steps in the EPAR. When this did not yield more consistent results, total and active review times in the annual reports were leading.

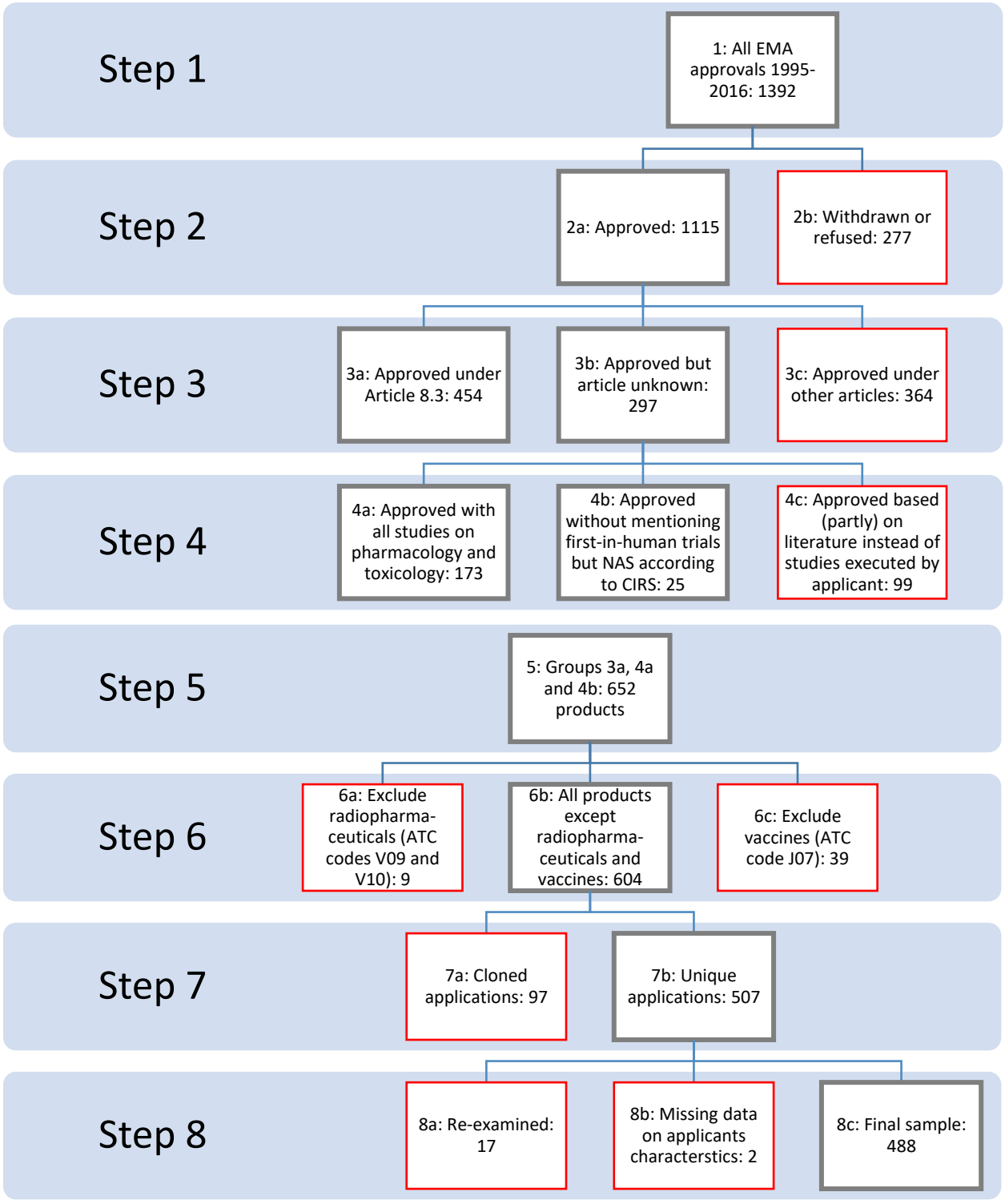


Figure 2: Sampling steps to arrive at the final sample of 488 approved products

Total review time can still be considered secondary endpoint, since clock stop time is likely depending on the gravity of the list of questions (LoQ), which the CHMP sends to the applicant after the initial assessment of the clinical data. It is likely that applicants which are expected to deliver high quality, based on the reputational judgement, receive less extensive LoQs. Having fewer questions to answer, results in shorter clock stop time and therefore clock stop time might to lesser extent also reflect the perceived organizational reputation by the regulator. Figure 3 visualizes the proportions of active review, clock stop and total review in comparison to the full review procedure of the EMA. This figure also emphasizes that CHMP opinion is given at day 210, but in practice this timeline is often deviated from and many submissions receive CHMP opinion before or after 210 days (Shah et al, 2013).

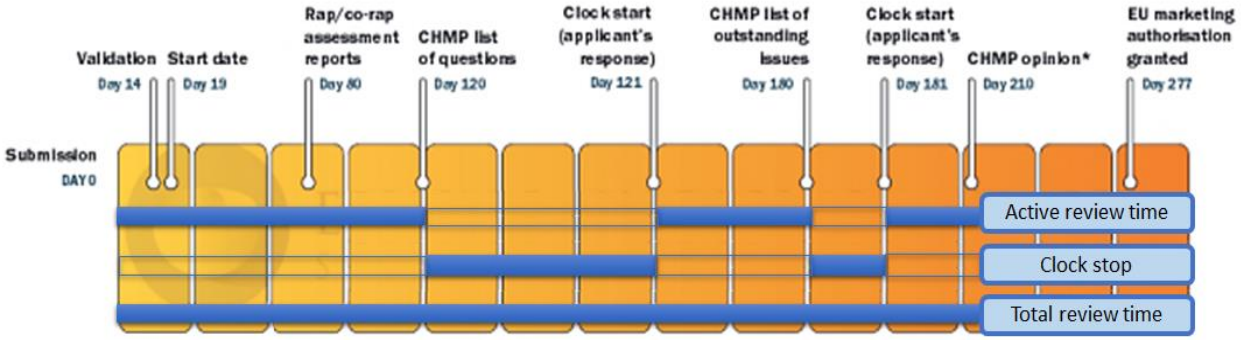


Figure 3: Buildup of review times for decision-making of market authorization at the EMA.
Adapted from: Beishon (2014)

3.4 Independent variables

The reputational judgement by the regulator depends on if applicants are known, known for something and their generalized favorability. Being known has been divided in past approvals and current applications, which reflect both the familiarity between regulator and applicant and emphasizes on the accumulated knowledge the applicant has about the procedure. Being known for something has been included by looking at quality issues and product discontinuities. Generalized favorability has been measured looking at general safety scandals and the degree of specialization of the applicant. The hypotheses about quality issues (H3), product discontinuities (H4) and safety scandals (H5) have been tested by two separate variables each.

3.4.1 Being known

Past approvals (H1). The number of past approvals has been measured by counting all products (including generics) of the applicant, that have been approved by the CHMP before the submission date of the product application. I assume that an applicant with many prior approvals is well-known by the CHMP-reviewers and the number of interactions is likely to speed up the review procedure. Because generic applications also allow interaction and information exchange between the CHMP and the applicant, excluding generic applications would rule out many interactions between regulator and applicant. I included procedures of approved generic products, which is under the assumption that there is no difference regarding applicant-regulator interaction between generic applications and applications under article 8.3. To calculate this measure, all approved products since the establishment of the FDA were included except for cloned applications (913 procedures). This ensured that companies that mostly focus on generic drugs and occasionally submit products under article 8.3, such as Teva, Intas, Mylan or Krka d.d. Novo Mesto would be regarded as having almost no experience when only including the approved products from the sample, despite their frequent interaction with the regulator and regulatory procedures.

Current applications (H2). The number of current applications is measured by the absolute number of ongoing applications, including generics that the applicant had at the time of the application

submission. The number of current applications is measured by a time-varying variable that has, similarly to the number of previous approvals, been calculated by including all previously approved products (excluding cloned applications). Similar to the number of previous approvals, this assumes that there is no difference in reputation-building between generic procedures and procedures under article 8.3. Olson (1997) argued that applicants that have more ongoing review procedures, possibly have more contact with the regulator which leads to both better information channels between applicant and regulator and produces reputational advantages since the applicant is already known among the reviewers. Stern (2017) said that the number of current applications reflects the expertise of the applicant in 'navigating the regulatory process' at that point in time.

3.4.2. Being known for something

Product referrals (H3). This variable measures the proportion of products of the applicant that has been subject to a referral procedure under article 20 of the Regulation (EC) 726/2004, before the CHMP opinion date of the product under assessment. After marketing authorization, the safety of new products is monitored throughout the product life cycle and reassessed if efficacy or safety problems occur (Bouvy et al., 2016). Several types of procedures exist to re-assess the product, but only the procedure under article 20 of the Regulation (EC) 726/2004 focuses on safety or quality issues of products that have been authorized through the centralized procedure (Bouvy et al., 2016). This variable measures the proportion of referrals regardless of the outcome, because when a product is reassessed by the CHMP through an article 20 procedure, doubts exists about whether the risk benefit ratio of the product is deemed positive enough, which harms the reputation of the applicant.

To compute this variable, a list of all referrals of centrally authorized products was composed by searching the publicly available EMA referral database. Data from the referral database was complemented with the data by Bouvy et al. (2016). A total of 85 centrally authorized product referrals were found within the chosen timeframe. For these products, the product name, referral opinion date by the CHMP, marketing authorization holder at the time of CHMP opinion were marked. Next, all referrals for cloned applications were removed since these are referrals about products with identical active substances and do not yield additional negative reputation, leaving a total of 56 referrals of unique INNs. The documents containing questions and answers about the reviews of these products were manually searched for the referral outcome and referrals were classified as (1) maintained, (2) obligated to some variation in product information, prescription or treatment, or (3) suspended.

Safety withdrawals (H3). Kim (2012) emphasized that the number of product withdrawals that an applicant has caused can be seen as a measure of reputation. In his analysis he included a count of the number of product withdrawals that an applicant has caused in the past. As only few safety withdrawals have occurred in Europe, I assume that having a negative reputation regarding safety withdrawals is long-lasting and therefore I use a binary variable that measures if the applicant had any products that have been withdrawn due to safety issues. To measure the effect that safety withdrawals has on the reputation of the applicant, a list containing all safety withdrawals in Europe between 1995 and 2016 was composed. The list was made by extracting all product withdrawals from the EMA website and searching the EMA public statement for the reason of withdrawal. The data from the EMA website were complemented by two scientific papers by Motola et al. (2006) and McNaughton et al. (2013), resulting in eleven unique safety withdrawals. The decision date of the European Commission in the EMA public statement was used as reference date, since some safety withdrawals were not imposed centrally, but rather non-renewed or withdrawn under the title of commercial reasons and therefore had no CHMP opinion date.

Voluntary withdrawals (H4). The effect of voluntary withdrawals is captured by the proportion of products, excluding vaccines, that has voluntarily been withdrawn by the applicant within five years after its initial marketing date. Taking this measure is under the assumption that the CHMP

incorporates all past approvals when making a judgement about the reputation of the applicant and that the MAH at the time of withdrawal and not the initial applicant is held responsible for the withdrawal. Also, this measure does not account for how often the drug is in fact still used. If the applicant decides to withdraw a product from the market because nobody is actually using this drug anymore because of a better alternative, it is likely to harm the reputation of the applicant to a lesser extent compared to a drug which is still being used, but withdrawn because the patient population is too small to have a commercially viable production. Despite these downsides, this measurement is the most accurate representation of the possible effect that voluntary withdrawals can have on the CHMP reputational judgements which can be derived from my data.

The list of product withdrawals, as described in the previous section, used to track all safety withdrawals was also used to compose a list of voluntary withdrawals. After excluding all safety withdrawals, the reason for withdrawal was limited to 'not marketed', 'commercial reasons' or 'non-renewal'. Since all three of these reasons are related to mismatching with a possible patient population or market or being outperformed by alternatives, all of these withdrawals were included as voluntary withdrawals. Withdrawals of cloned products, while at least one other product with the same active substance and MAH remained available, were excluded. For the list of all voluntary withdrawals between 1995 and 2017, which totaled 115 products, the decision date of the European Commission or the MAH at the time of withdrawal were extracted from the EMA public statements, available in the online EPAR.

Failed applications (H4). This variable is measured by the proportion of applications that were refused by the CHMP, or withdrawn by the applicant before CHMP opinion, as compared to the total number of applications by that applicant. The effect of failed applications was operationalized in two ways. Initially the absolute number of failures was included, but since applicants with more applications logically also face higher chances of failing and therefore the second operationalization was to compute the failure rate as a percentage of all submitted applications (approved, withdrawn and refused). Using the failure rate has the underlying assumptions that the CHMP is sensitive to the relative amount of successful products an applicant has produced in the past and that its memory lasts, even when members of the CHMP are replaced. Using the failure rate is also under the assumption that once organizations change, such as happens for mergers or name changes, the CHMP has no cognitive memory of the former organization, meaning that the failure rate restarts at zero.

To be able to include these effects in the analyses, a list of all product refusals and withdrawals in the used timeframe was composed. Application refusals and withdrawals were provided by Hoekman, simultaneously with the list of product approvals. This includes only withdrawn applications and refusals from 2004 onwards, since the EMA has not made failed applications publicly available before that year. The CHMP refusal dates were obtained from the documents with procedural steps, stating the date when the CHMP issued a negative opinion. For withdrawn applications, the withdrawal date stated at the EMA website was taken, since product withdrawals do not require an opinion by the CHMP.

3.4.3 Generalized favorability

Referrals (H5). This variable measured if a product by any firm in the industry was suspended after a risk-benefit reassessment under article 20 of the Regulation (EC) 726/2004, in the 365 days prior to the CHMP opinion date. Kim (2012) used a binary variable for review procedures with safety withdrawals in the year prior to the application to account for its reputational effects and found that these have a significant negative influence on review times at the FDA. Using a binary variable indicates that the effect that the number of product suspensions has on the generalized favorability of applicants is noncumulative, meaning that the CHMP will not be more cautious when multiple product suspensions happened in the prior year. This variable has been constructed by using the list with all product referrals, as constructed for the variable "product referrals" (H3). The list was list was limited

to only suspended products, since product suspensions indicate serious safety issues and are therefore more likely to affect generalized favorability than products that were maintained or subject to small changes in prescription or treatment. This resulted in eleven suspended products. The referral date was taken and using RStudio, the number of product suspensions in the year prior to the CHMP opinion was calculated.

Time since last safety withdrawal (H5). The effect that safety withdrawals have on organizational reputation is measured by a categorical variable that categorizes the products according to the amount of time between the last safety withdrawal and the moment of CHMP opinion. To include the possible effects of safety withdrawals for the CHMP, primarily a possible operationalization was to count the number of days between product submission and the latest safety withdrawal. But since the first safety withdrawal was from 2000, this resulted in almost 100 missing values. To overcome this problem a continuous and a categorical variable were composed using the list with all safety withdrawals, as composed for the variable "safety withdrawals" (H3). The five categories that were used to indicate the time since the last safety withdrawal are: (1) no prior safety withdrawals, (2) latest safety withdrawal within 183 prior to CHMP opinion, (3) latest safety withdrawal within 184 to 365 days before CHMP opinion, (4) latest safety withdrawal within 366 to 730 days before CHMP opinion and lastly (5) the latest safety withdrawal more than 730 days before CHMP opinion. Using this categorical variable assumes that the memory that the CHMP has about these safety issues gradually fades away over time.

Specialization (H6). The degree of specialization is measured by the percentage of sales from pharmaceuticals as compared to the total sales of the applicant. The degree of specialization is computed by dividing the sales from pharmaceutical products in the year prior to product approval by the total sales from the company in the year prior to product approval. Data gathering was done by using Scrip Pharma Intelligence, both the books with published financial data from 1994 and 1995 and online data from the years thereafter. Missing data was gathered by searching the annual reports and financial tax statements (US firms). If there was sufficient reason to believe that the no sales occurred prior to the MAA (e.g. first financial statements from the year after MAA approval, no existence of company news before MAA), sales were set to zero. For the MAAs of SonoVue and Axura, submitted by applicants Bracco and Merz Pharma, no reliable data on sales could be found and therefore these MAAs were excluded from the analyses. This has already been visualized in figure 2, reducing the number of procedures included from 490 to 488.

3.5 Control variables

Product uncertainty is measured by looking at product complexity, the size of the evidence base at the time of the CHMP opinion and whether the product has been authorized by a comparable regulatory body, which in this research is the FDA. Status is measured by the average number of citations on scientific publications, which emphasizes the ranked position of the applicant when looking at the prerequisites for producing quality products (Rindova et al., 2006). Other controls are sales of pharmaceutical products by the applicant, age, origin and whether the MAA was submitted for accelerated approval.

3.5.1 Product uncertainty

Complexity. To grasp product complexity, the number of words to describe the pharmacological properties and chemical structures in the product label has been measured. This is an imitation of Moffit (2010), who argued that if less knowledge was available about a product at the time of approval, a more extensive description of product features is required in its product label. To retrieve description length, an automated word count has been performed using RStudio, counting all words in the EPAR (excluding numbers and punctuation) in sections 4.2, 5.1 and 5.2 of the Summary of Product Characteristics (SPC) issued at the time of marketing authorization. Section 4.2 describes drug posology

and method of administration, whereas sections 5.1 and 5.2 respectively describe the pharmacodynamics and pharmacokinetic properties of the product.

Evidence base. To measure the evidence base of the product under assessment, the size of the clinical program at the time of the product application stated in the product EPAR is included. Van Duijnhoven et al. (2013) argue that available knowledge about product efficacy and safety to base the product assessment on, is largely dependent on the amount of evidence available from the clinical program. This measurement accounts for the uncertainty the CHMP faces when determining the benefit-risk balance of a new product, because if limited information is available about possible adverse side effects and the product efficacy, it is harder for the CHMP to determine the risk-benefit balance (van Duijnhoven et al., 2013). To collect the data on clinical program size, the publicly available data from van Duijnhoven et al. (2013) has been used, complemented by the total number of healthy volunteers and patients that have been exposed to the drug prior to marketing authorization. This data was manually extracted from the product EPARs (section 'Clinical aspects'). Since prevalence largely differences across disease (e.g. orphan diseases) and finding patients that are willing to participate in clinical trials is most likely easier when prevalence is high, this measurement simultaneously controls for disease presence. Nevertheless, treatment length and intensity varies per disease, which is not accounted for, since only the absolute number of unique subjects (healthy volunteers and patients exposed to the drug) that were engaged in safety and efficacy trials is included.

Approval history. The approval history of the product under assessment is measured by a categorical variable that categorizes products according to the time that has passed between FDA approval and CHMP opinion. Many new drug applications are filed primarily in the US and additionally, the FDA generally is quicker to allow new products market entry (Beishon, 2014; Roberts et al., 2011). Therefore, products are regularly already available in the US when an application for the same product is filed at the EMA. The EMA and FDA perform similar risk-benefit analyses and information is increasingly shared between these organizations, but opinions can still diverge because of different evaluations of (clinical) endpoints (Roberts et al., 2011; Tafuri et al., 2014). Because risk-benefit analyses are similar, FDA approval can be regarded as an indicator of product uncertainty, since receiving marketing authorization by the FDA shows that the product is regarded as having a positive benefit-risk ratio. Also, if more time expired between FDA approval and application at the EMA, additional clinical data has been collected and more knowledge surrounding efficacy and possible side-effects is present.

This means that the longer the product already has been approved by the FDA, the more well-considered the CHMP's opinion can be. Hence, the categories are (1) no FDA approval prior to CHMP opinion, (2) FDA approval within 183 days before CHMP opinion, (3) FDA approval within 184-365 days before CHMP opinion and (4) FDA approval more than 365 days before CHMP opinion. A categorical variable is preferred over a binary variable, since it is more distinctive in the time that has passed since FDA approval. The FDA approval dates were kindly provided by Hoekman, following earlier research by Hoekman and Boon (2019).

3.5.2 Status

Average citations on publications. Here, the average number of citations on Web of Science on the scientific output of the applicant in all years prior to the MAA submission is used as status measurement. According to Polidoro and Theeke (2012) scientific journals request certain standards regarding quality and integrity, which constitute an important form of certification that may help third parties to assess innovations, such as the CHMP in its efforts to assess the quality of the applicant. The number of citations is generally accepted as indicator of the quality of publications and has been used prior by Hu et al. (2015) to indicate social status. Whilst Kim (2012) argued that it is preferred to derive status from patent counts, I argue that publications are a more accurate measurement here, since in

drug development, scientific publications occur after approval and thus do not distort the data by including publications about the product under assessment. To find the number of citations on scientific publications, searches were performed using a time period limited to the year prior to the submission year, including only English written scientific articles and searching for the name of the organization using the search terms provided in Appendix A.

3.5.3 Other controls

Accelerated assessment. There are several special pathways an applicant can apply for when submitting a new product application. Since some of these pathways are designed to speed up the regulatory process it is important to include these as control variables (Kim, 2012; Olson, 1997; Stern, 2017). The pathway included here is whether the product has been submitted for accelerated assessment (AA). Other pathways (i.e. conditional market approval (CA) and approval under exceptional circumstances (EC)) are excluded here because these are not meant to lead to faster reviews by the CHMP, but only to provide marketing authorization with fewer clinical data. The data regarding an application for expedited pathways was provided by Hoekman. Applying for AA is not equal to receiving an AA, meaning that it is a necessary condition but not automatically sufficient.

Size. The size of the applicant is measured by looking at the pharmaceutical sales (in million $) of the applicant in the year prior to the CHMP opinion date, since the year of CHMP opinion likely already includes sales of the product under investigation. There is a need to control for the size of the applicant, since this already has been identified as an indicator of quality (Olson, 1997; Kim, 2012). Both data on the total sales and the assets that an applicant possessed was gathered through the SCRIP database (see description at variable specialization), annual reports and tax statements. The value of assets and pharmaceutical sales had comparable correlations with other independent variables, but including assets resulted in more missing data points. Therefore, the size is measured by the size of the company's pharmaceutical sales.

Age. The age of the applicant is calculated by the difference between the founding year of the firm and the year that the application was submitted. This does not per se define the size of the company, but it is important to control for, especially since reputation is said to be built up by the past of a company (Bitektine, 2011). The founding year of the company has been extracted from the Bloomberg website with company profiles, using the applicant name as search term. When doubt existed about the data, it was checked using online company database Crunchbase. This would happen for applicants with a founding date after the submission date of their first product, two company profiles with different founding years or no founding year at all. The age of the company was calculated by subtracting the founding year by the year of submitting the application.

Origin. Since Olson (1997) found that applicants located in the US received significantly faster approvals than foreign firms at the FDA, it is important to control for the origin of the applicant. A possible mechanism that might cause local applicants to receive faster approvals is that they are often geographically more closely located to the regulator, which could strengthen information channels and personal connection with the applicant. However, I found no scientific theoretical evidence that supports the findings by Olson (1997) for the European context. Data on the origins of the applicants was provided by Hoekman, who had made an overview of companies including their origin, for earlier research (Hoekman & Boon, 2019)

Table 5 summarizes all variables as described in section 3.3, 3.4 and 3.5 and includes the measurement and data sources used to obtain the data.

Table 5: Operationalization of variables and used data source

Dependent variable			
Indicator	**Type**	**Measurement**	**Data source**
Active review time	Ratio	Total review time in days minus 'clock stop' time	EMA annual reports annex 'CHMP opinion in year x on medicinal products for human use'
Independent variables			
Indicator	**Type**	**Measurement**	**Data source**
Product uncertainty - Complexity (ln)	Ratio	The number of words used to describe sections 4.2, 5.1 and 5.2 of the SPC	Automatically extracted from SPC at time of approval (SPCs already gathered for earlier research by Boon & Hoekman, 2019)
Product uncertainty - Evidence base (ln)	Ratio	The total number of healthy volunteers and patients that were engaged in clinical safety studies indicates in the SPC	EPAR section 2.4, or otherwise the accumulated total number of unique participants in the clinical studies.
Product uncertainty - FDA approval	Categorical	Days between FDA approval and CHMP opinion: 1 = Not approved by FDA 2 = FDA approval <183 before CHMP opinion 3 = FDA approval 184 to 365 days before CHMP opinion 4 = FDA approval >365 days before CHMP opinion	FDA approval dates were already gathered for earlier research by Boon & Hoekman (2019)
Being known - Past approvals (ln)	Ratio	The number of previous approvals by that applicant (including generic applications)	Application data was already gathered for earlier research by Boon & Hoekman (2019)
Being known - Current applications	Ratio	The number of products under assessment by the applicant during the review period (including generic applications)	Application data was already gathered for earlier research by Boon & Hoekman (2019)
Being known for - Product referrals	Ratio	The percentage of products produced by the applicant that has been subject to a risk-benefit reassessment. Percentage of all previously approved products (including generics)	Extracting the referral database from the EMA website, complemented by manually checking CHMP minutes and meetings summaries and referral data kindly provided by Bouvy et al. (2016).
Being known for - Safety withdrawals	Binary	Has the applicant previously produced a product that has been withdrawn for safety reasons (1) or not (0)	Safety withdrawals were gathered by extracting all withdrawals from the EMA website and manually checking them for safety reasons. This approach was checked and complemented by the papers by Motola et al. (2006) and McNaughton et al. (2013).
Being known for - Voluntary withdrawals	Ratio	The percentage of products produced by the applicant that has voluntarily been withdrawn from the market within 5 years (excluding generics), as percentage of total number of approvals	All withdrawals have been extracted from the EMA website and classified according to reason of withdrawal. The reasons were "non-renewal', 'commercial reasons', 'not marketed' or 'safety or quality issues'. All products that were not withdrawn due to safety issues were classified as voluntary withdrawals.
Being known for - Failed applications	Ratio	Percentage of products refused or withdrawn during the review process compared to the total number of submitted applications	Refusals and applications were already gathered for earlier research by Boon & Hoekman (2019). Opinion and withdrawal dates were distilled from the EPARs and EMA website
Favorability – Product suspension	Binary	A binary variable that measured if any product suspension under Regulation (EC) 726/2004 article 20 happened in the year prior to CHMP opinion.	Extracting the referral database from the EMA website, complemented by manually checking CHMP minutes and meetings summaries and referral data kindly provided by Bouvy et al. (2016).

Favorability – Time since last withdrawal	Categorical	Days between the last safety withdrawal for any firm in the industry and CHMP opinion of the MAA under assessment: 1 = No prior safety withdrawal 2 = Safety withdrawal within 183 days prior to CHMP opinion 3 = Safety withdrawal 184 to 365 days before CHMP opinion 4 = Safety withdrawal 366 to 730 days before CHMP opinion 5 = Safety withdrawal >730 days before CHMP opinion	Safety withdrawals were gathered by extracting all withdrawals from the EMA website and manually checking them for safety reasons. This approach was checked and complemented by the papers by Motola et al. (2006) and McNaughton et al. (2013).
Favorability – Specialization	Ratio	Pharmaceutical sales as a percentage of total sales	SCRIP yearly top 150 complemented with data from annual reports
Status (ln)	Ratio	The average number of citations on all scientific publications published by any author employed by the applicant, in the year prior to the MAA submission	WoS search for organization name. Limiting search to only English scientific articles from 1900 until the year prior to the application submission. Create citation report and extract average citations per publication.
Applied for AA*	Binary	1 = Yes 0 = No	Data gathered for earlier research by Boon & Hoekman (2019)
Size (ln)	Ratio	Pharma sales in million US$ in year prior to application submission	SCRIP yearly top 150 complemented with data from annual reports
Age (ln)	Ratio	Year of drug application submission minus founding year of company	Founding year of applicant is retrieved from the Bloomberg website with company profiles
Origin	Binary	Whether the applicants headquarters are located in Europe (1) or not (0)	Data gathered for earlier research by Boon & Hoekman (2019)

** AA = Accelerated Assessment*

3.6 Data analysis

3.6.1 Negative binomial regression

The dependent variable in this research is a count variable and does not follow a normal symmetrical distribution, because its distribution is both negatively skewed (-2.038) and has a kurtosis that is leptokurtic (5.864). Because the distribution of the dependent variable violates the assumption of normality, an ordinary least square regression is not suited. The active review time measures the number of days since the start of the review procedure and thus can be considered a count variable, as it cannot take any negative values, which advocates the use of count models. Using these models, i.e. (quasi) Poisson and negative binomial models have been used before to analyze the effects of a set of predictors on count variables (e.g. Polidoro & Theeke, 2012). Count models estimate what predictors generate the outcomes that are most closely matching the values that are observed in the sample, maximizing the likelihood of obtaining the sample values.

The normal Poisson model has the underlying assumption of equidispersion, meaning that the sample mean (μ) is equal to the variance which is the squared standard deviation (σ^2) (Ver Hoef & Boveng, 2007). In this sample, the assumption of equidispersion is violated, because the data is over dispersed (see formula 1) which means that the variance in the sample is significantly ($p<0.001$) greater than the theoretical variance of one, which is used in the Poisson models. Using a Poisson in this case would result in small standard errors and therefore the Poisson is likely to find significance levels that are too generous. In sum, using a Poisson model to analyze sample data which is overdispersed, can result in finding effects that are not really present.

$$Formula\ 1\text{: Dispersion} = \frac{\sigma^2}{\mu} = \frac{22.37^2}{196} = 2.55$$

To overcome the problem of overdispersion, several statistical models have been developed that provide a solution such as the quasi-Poisson model and the negative binomial model (ver Hoef & Boveng, 2007). Both methods calculate a dispersion parameter from the data and does not assume the dispersion to be fixed at one. The difference between these models is that the quasi Poisson method assumes a linear relation between σ^2 and the μ, whereas the negative binomial reasons from a quadratic relationship and uses an additional parameter to adjust for variance independently from μ. These differences results in different estimated weights, quasi Poisson weighing proportionally to the mean, while negative binomial assigns little weights to both very small and large means (ver Hoef & Boveng, 2007).

Both the quasi-Poisson and the negative binomial model have been executed in this research to calculate the size of predictors on the active review time by the CHMP. Quasi-Poisson and negative binomial models have been executed multiple times using different sets of predictors, as summarized in table 6 below. The 'control' model only includes control variables. Models 1 until 9 only sequentially include one reputation variable to account for its specific effects. Firstly, I test the effects of the predictors that are used as controls and subsequently for variables that proxy being known (1-2), being known for quality issues (3-4) or product discontinuities (5-6) and lastly indicators that measure generalized favorability (7-9). After including the predictors separately, their effects are also estimated by combining the predictors within each dimension.

Table 6: Summary of regression models used for data analysis

	Control	Being known			Being known for something					Favorability				All
Variable		1	2	1-2	3	4	5	6	3-6	7	8	9	7-9	
(U) Complexity	x	x	x	x	x	x	x	x	x	x	x	x	x	x
(U) Evidence base	x	x	x	x	x	x	x	x	x	x	x	x	x	x
(U) Approval history	x	x	x	x	x	x	x	x	x	x	x	x	x	x
(BK) Past approvals		x		x										x
(BK) Current applications			x	x										x
(BKS) Product referrals					x				x					x
(BKS) Safety withdrawals						x			x					x
(BKS) Voluntary withdrawals							x		x					x
(BKS) Failed applications								x	x					x
(GF) Product suspension										x			x	x
(GF) Time since withdrawal											x		x	x
(GF) Specialization												x	x	x
Status	x	x	x	x	x	x	x	x	x	x	x	x	x	x
AA	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Size	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Age	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Origin	x	x	x	x	x	x	x	x	x	x	x	x	x	x

(U) = variables measuring product uncertainty, (BK) = variables measuring being known, (BKS) = variables measuring being known for something; (GF) = variables measuring generalized favorability

3.6.2 Cox proportional hazard models

After having performed the main analyses using quasi Poisson and negative binomial estimation methods, additionally the data was analyzed using the Cox proportional hazard models. The reason to use the latter as a separate analysis is because count models assume a Poisson distribution of non-negative integers of the dependent variable. The active review time is indeed measured by a non-negative integer, but does not follow a clean Poisson distribution, which advocates for a second analysis. Hazard models have been used as an alternative for generalized linear models, even when there was no need for censoring and do not require the data to be Poisson distributed (Basu et al., 2004). These methods also have been used in the similar work by Carpenter et al. (2010) and Kim

(2012), who used hazard models as main method of analysis to explain variation in review times at the FDA.

I have performed two separate regressions using the Cox proportional hazard method. The primary regression includes all observations as having had the event (receiving approval), so no distinction between observations was made regarding their predicted survival. Secondly, I performed a regression using a cutoff at 200 days, were all MAAs approved before day 200 were scored as having had the event (approval) and all MAAs that were approved after 200 days of active review received score zero, indicating they had not had the event. The reasons for using the 200 day mark as cutoff is explained in the remainder of this section, but first I emphasize on the proportional hazard assumption and its implications.

Proportional hazard assumption. The proportional hazard assumption implies that ratio of hazard for independent observations is constant over time. In other words, it assumes that the effect of predictors on the outcome is constant over time (Schober & Vetter, 2018). To investigate the hazard proportion in my data, a Kaplan Meier survival curve has been plotted, which is visualized in figure 4. Because the line is not decreasing linearly, there is a suspected violation of the hazard assumption and thus a Schoenfeld's residual test was performed. This test determines whether the residuals of the β-coefficient of predictors significantly differ over time, testing goodness-of-fit using X^2 (Schoenfeld, 1982). The X^2 scores and p-values of the individual predictors are given in table 7 below. If the proportional hazard regressions are executed with predictors that vary over time, the effects found for these predictors will merely be an average of the effects over the whole timespan.

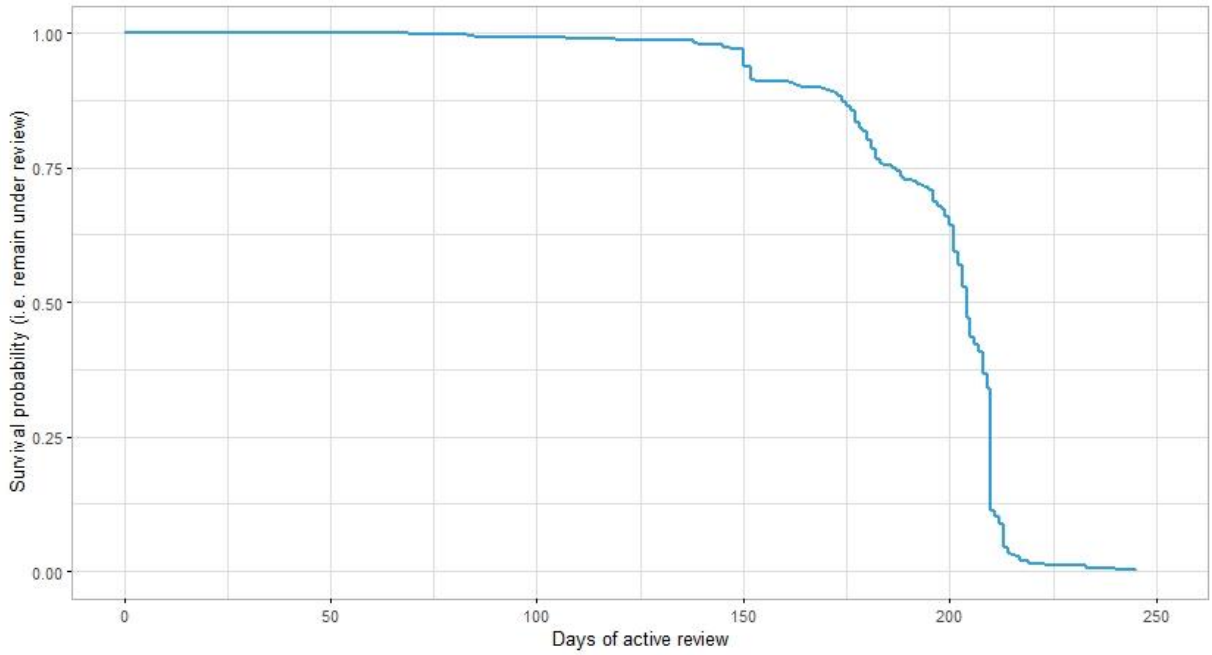


Figure 4: Kaplan Meier plot estimating survival probability

From Schoenfeld's test it appears that the β-coefficient of the variables applying for accelerated assessment, and two of the variables that measure generalized favorability (the time since the last safety withdrawal and the number of product suspensions in the prior year) are significantly time dependent ($p<0.001$). The plotted residuals of variables that significantly differed over time are provided in appendix B. The variable measuring if was applied for accelerated assessment is highly time dependent because products that have been granted accelerated assessment are reviewed preferably within 150 days, but ultimately in 180 days, meaning that the effect size should dissolve to zero after day 180. The variables measuring the effect of general safety issues (product referrals and

safety withdrawals) also violate the proportional hazard assumption. They have a predicted negative effect for MAAs that have an active review time of less than 210, which becomes positive closer to the 210 days mark.

Table 7: Outcome of Schoenfeld residuals test

Variable	X^2 scores active review time	X^2 scores using cutoff at 200 days	X^2 scores total review time as DV
(U) Product complexity	0.708	1.141	0.778
(U) Evidence base	0.162	0.147	3.851
(U) <183 days since FDA approval	2.529	0.103	0.584
(U) 183 to 364 days since FDA approval	1.869	2.829	8.413**
(U) >365 days since FDA approval	0.000	0.001	1.902
(BK) Past approvals	1.526	0.088	0.155
(BK) Current applications	4.650	1.828	0.101
(BKS) Safety withdrawals	3.509	0.656	0.176
(BKS) Product referrals	0.039	0.153	0.128
(BKS) Voluntary product withdrawals	0.127	1.066	0.007
(BKS) Failed applications	1.867	0.030	0.503
(GF) Number of product suspensions prior year	12.449***	1.238	2.133
(GF) <183 days since last safety withdrawal	17.319***	3.986*	2.791
(GF) 183 to 364 days since last safety withdrawal	17.748***	2.003	0.941
(GF) 365 to 730 days since last safety withdrawal	13.644***	8.107**	4.601*
(GF) >730 days since last safety withdrawal	30.233***	14.375***	3.006
(GF) Specialization	0.202	0.592	1.704
Status	0.594	0.288	0.267
Age	0.157	0.468	1.365
Sales	1.732	2.936	0.001
Applied for AA	46.018***	21.536***	9.481**
Origin	1.227	0.621	0.361

*Note: *=p<0.05;**=p<0.01;***=p<0.001; DV = Dependent variable, U = Variables measuring product uncertainty, BK = variables measuring "being known"; BKS = variables measuring "being known for something"; GF = variables measuring "generalized favorability"*

Proportional hazard using a cutoff. To account for the differences in effect sizes over time, a second analysis has been performed using a cutoff at day 200. This means that all products approved before 200 days of active review time were given score 1 in an additional event-variable. The cutoff point of 200 days was chosen because as figure 5 visualizes, the change in survival rate before and after that day are quite distinct. Between days 150 and 200, the survival probability decreases with 29.39%, whilst between days 200 and 213 it is 55.51%. This is most likely due to the pre-arranged timeline the CHMP uses to review products, having the ability to grant faster reviews, but avoiding crossing the 210 days mark. The reason some products are reviewed in 211, 212 or 213 days is because the CHMP meetings are every month, during a three-days gathering. If a review was started at day one of the gathering, and several meetings later an opinion has been filed at the last day, the 210 days is likely to be extended by up to three days. By using the cutoff of 200 days, the estimated coefficients for the predictors show their impact on the hazard ratio of being approved before the 200 days mark. Schoenfeld's residual tests have been performed for the regression data with the cutoff at 200 days and given in table 7. It appears that using a cutoff date decreases the number of variables that have significant differences in effect size over time.

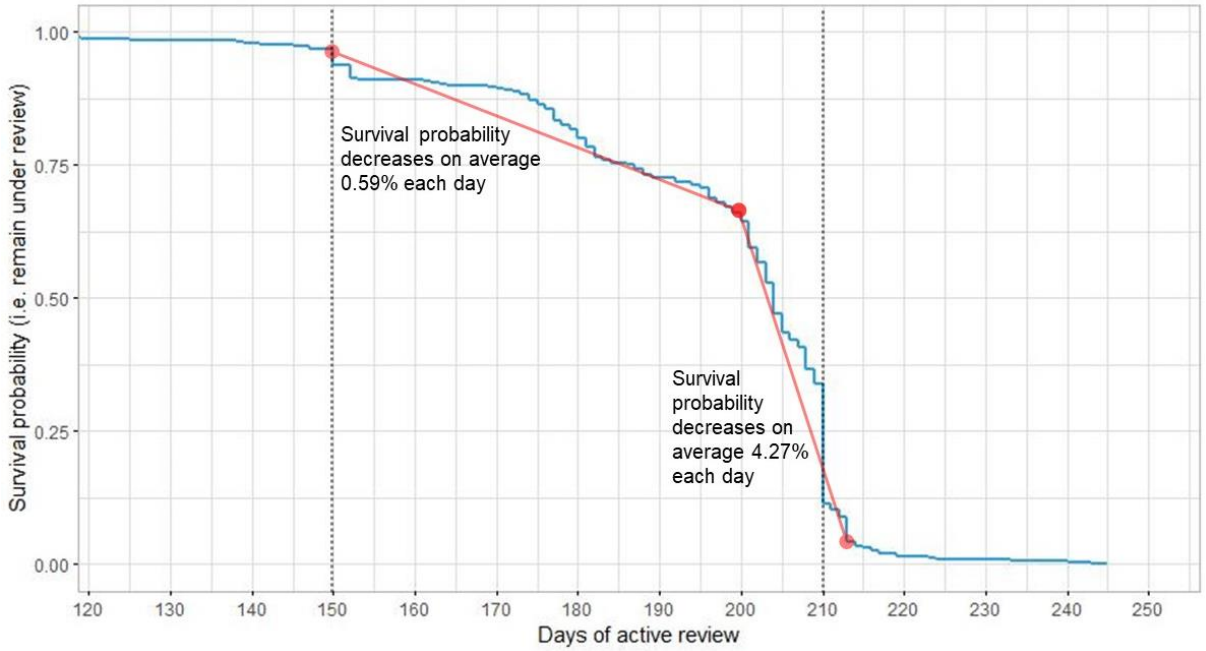


Figure 5: Kaplan Meier plot with estimated differences in survival probability

3.6.3. Robustness checks

Total review time as dependent variable. To verify the findings that were obtained via the negative binomial and Cox regression, these regressions were also executed using the total review time as dependent variable. As already argued in section 3.3, the total review time might also reflect the organizational judgement by the CHMP, as the clock stop time dependents on the severity of the list of questions the CHMP has composed. Table 7 includes the Schoenfeld's tests for the predictor variables using the total review time as dependent variable and has less problems with the proportional hazards assumption than using active review time as dependent variable with or without using a cutoff at day 200. Since violation of the proportional hazard assumption happens to lesser extent and no clear cutoff can be indicated in the distribution of the total review time, the Cox regression has been performed with no cutoff date.

Correct for autocorrelation. Because the unit of analysis in this research is the organization that submitted a product for review, the dataset contains multiple observations that are likely subject to autocorrelation. To determine if autocorrelation exists between observations and introducing clustered standard errors is indeed necessary, the Durbin-Watson test is most commonly used (White, 1992). Hence, I executed the Durbin-Watson test grouping observations by the applicant and found that autocorrelation is significantly greater than zero ($p<0.001$), which implies that autocorrelation exists. This justifies introducing clustered standard errors for the applying firm, which should adjust for autocorrelation and thus account for unobserved fixed firm effects (Thompson, 2011; Wooldridge, 2007). The clustered standard errors for the negative binomial regressions have been computed by using RStudio using Mackinnon and White's (1985) standard errors and the applicant as cluster, which should lead to more correct estimates of the standard errors (Petersen, 2008). For the proportional hazard method Therneau et al. (2019) have specified that clustered standard errors are only required when observations have multiple events, as is not the case in either one of the proportional hazard regressions. Hence, clustering of standard errors by applicant has not been performed for the proportional hazard regressions.

Introducing dummies for therapeutic classes. Olson (1997) found that the therapeutic class of the disease is linked to the time of active review, since some classes require more complex medications. To account for this possible effect, an univariate regression has been executed using dummies for each

therapeutic area, based on the main group of the products ATC code. Thereafter, an additional negative binomial regression has been executed to test for the differences amongst the therapeutic classes. If Olson's (1997) argument is valid and the complexity of each therapeutic class is not grasped by the product complexity variable, some of the therapeutic classes should have significant estimates after introducing the dummies.

3.7 Quality indicators

3.7.1 Reliability

To ensure high reliability of this research, I used highly cited scientific literature, acknowledged in several scientific fields, as theoretical background to develop my analyzing framework. To allow replication of this research, the performed proceedings to gather and analyze the data have been extensively described in chapter 3 and primarily secondary data that is freely available at the EMA website or annual reports, has been used. Using secondary data published by the EMA ensures high consistency of data since the EMA publicly discloses all drug applications and market authorization decisions yearly in its annual reports. Additionally, data operationalization has been described in section 3.3, which allows for reproduction of the performed data analyses.

By looking at previous operationalizations of reputation in the pharmaceutical industry, indicators were constructed that fit the specific context of the pharmaceutical industry. Additionally, the three dimensions of reputation have been measured by multiple indicators each, which allows for more specific interpretation of the results. Control variables have been used that were applied in various closely related empirical articles which focused on the regulatory process in the pharmaceutical market. To check consistency of outcomes, several statistical models (quasi-Poisson, negative binomial and proportional hazard models) have been used to estimate the effects of predictors.

3.7.2 Internal validity

Internal validity was pursued by measuring all independent variables prior to the approval decision and including control variables that account for possible variation in review time, but that are not the focus of this research. Firstly, gathering data for predictor variables prior to the approval dates disallows wrong causality. Second, by controlling for other variables that influence active review times such as expedited pathways, status, product and firm characteristics, I tried to include a variety of factors that might account for variance in active review time. This allows to make statements about the influence of reputation while not neglecting other possible predictors. This research only includes approved new drugs, and therefore it is impossible to make statements about the influence of reputation on the speed of review times of rejected applications.

3.7.3 External validity

The results found in this research are specific for the pharmaceutical industry, but are possibly also applicable for other industries where regulatory agencies monitor firm behavior and control the quality and market access of new products. These could for example be food processing and production or the banking industry, but also highly innovative industries that currently lack regulation such as for example the implementation of artificial intelligence in diagnostics of diseases. However, one should be careful with generalization, since regulatory duties and powers differ largely across industries and countries.

4. Results

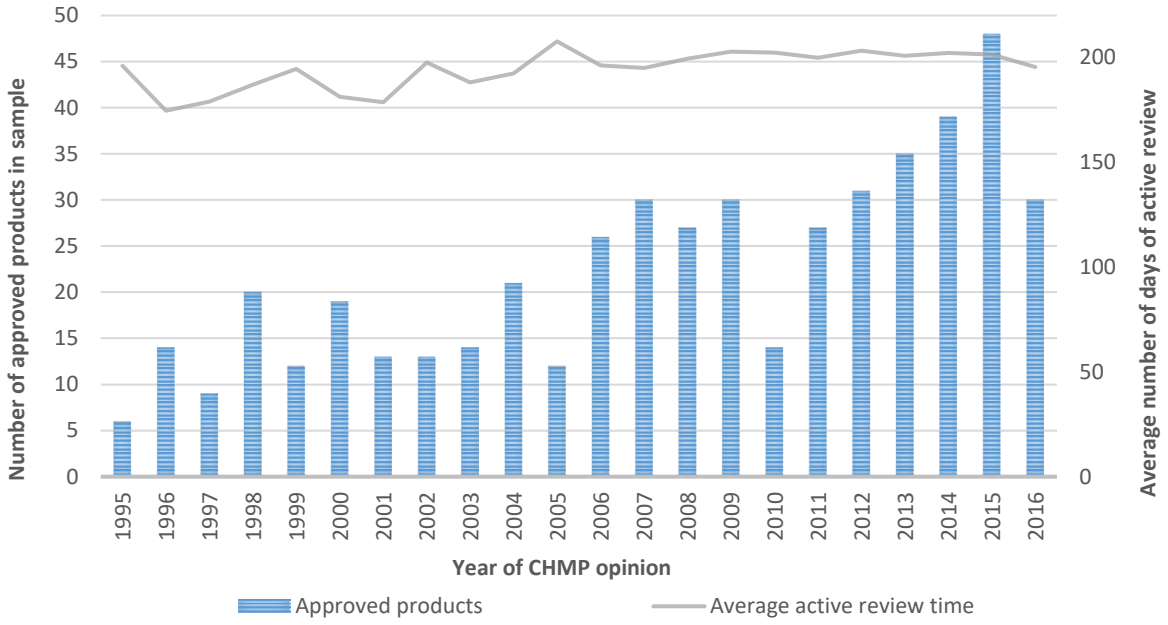
This chapter is devoted to describing the sample characteristics and gives the results of the statistical tests. In the first section the data structure has been described (4.1). The second section contains the results of the count models, focusing on the negative binomial regressions and the proportional hazard models (4.2). Lastly, several checks have been done to demonstrate robustness of the findings (4.3).

4.1 Data description

4.1.1 Dependent variable

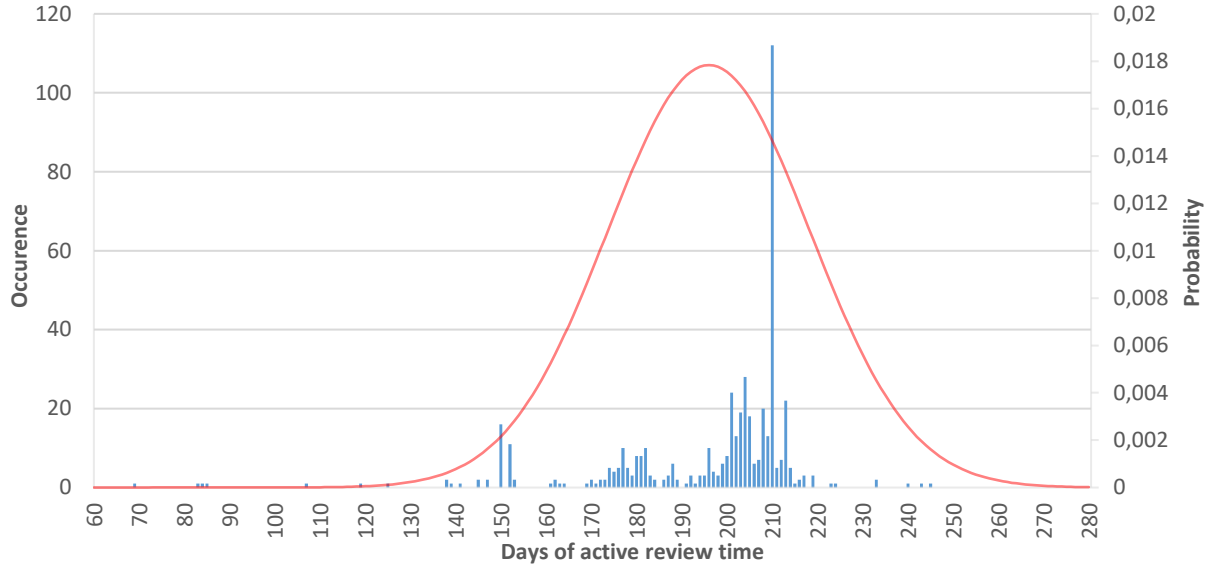
Active review time. The sample consists of 488 new drugs that were submitted and subsequently approved without re-examination by the EMA during between January first 1995 and the 31st of December 2016. The yearly average number of approved NMEs in this sample amounts 22, with a maximum of 48 in 2015 and a minimum of six in 1995 as shown in graph 1 below. The total review time averaged 357 days, of which on average 196 days active review time and 161 days clock stop. In 1996 the average active review time was shortest, counting 175 days. In 2005 average active review times was the longest with 208 days. Since 2005 the average active review time has fluctuated around the 200 days mark.

Graph 1: Yearly number of approved products and average active review time in sample



Graph 2 shows the sample distribution and a normal shaped distribution curve, which shows that the data is not normally distributed. In the sample a total of 111 approvals received CHMP opinion after exactly 210 days as is formally required (Shah et al., 2013). Of the other 377 approvals, respectively 322 and 55 procedures were completed faster and slower than 210 days of active review. In total, only 7 products (1.4%) took longer to review than 220 days, of which Eli Lilly's Humalog® had the longest active review time with 245 days. Abbott's Norvir® holds the shortest total and active review time with 69 days of active time and no clock stop time. The figure also shows a smaller peak around 150 days of active time, which is possibly due to the decision-making timeline and the agreements to assess products that are being reviewed under the expedited pathway of accelerated approval after 150 days. The 488 approved products in the sample were submitted by 162 different applicants, averaging 3.01 approvals per applicant. A total of 97 firms received approval for only one new drug within the given timeframe.

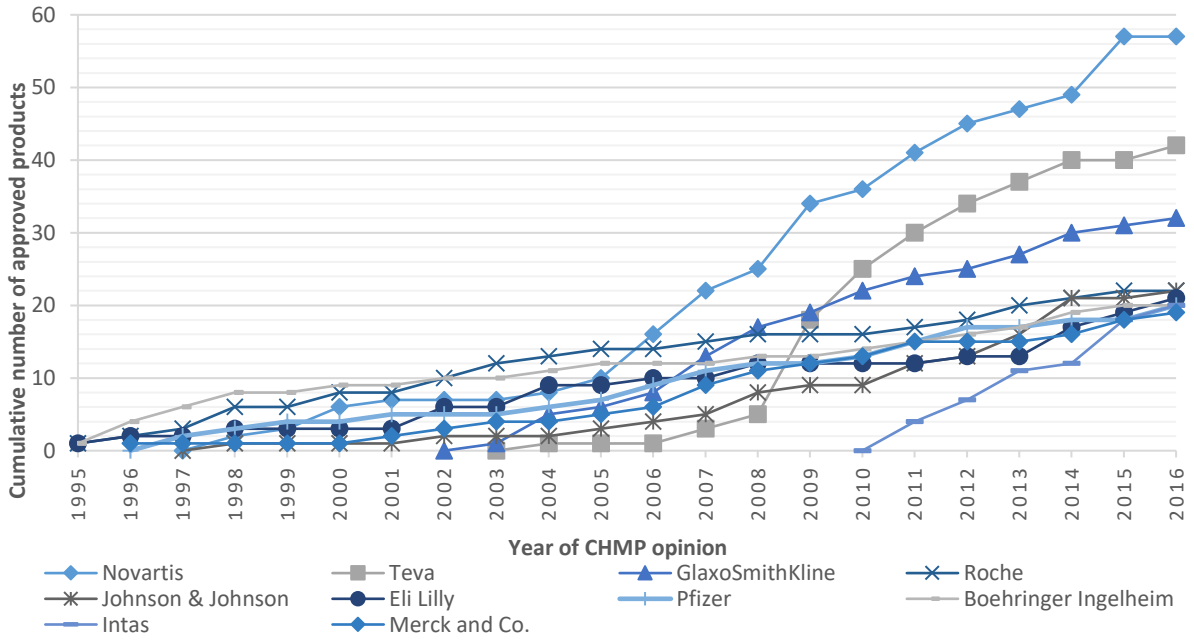
Graph 2: Density plot of active review time and normal distribution plot



4.1.2. Being known

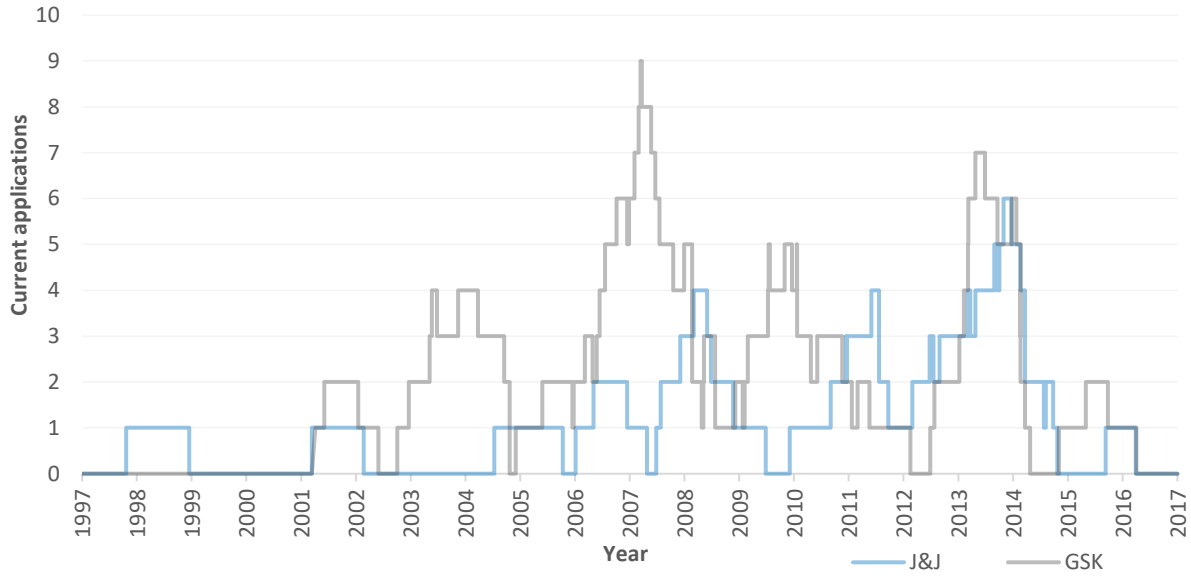
Previous approvals. Graph 3 shows the ten companies with most past approvals from the establishment of the EMA until 2017. The company with most past approvals in the sample is Novartis, having had 49 products approved before the submission of the application for Entresto® in 2015. Novartis is followed by Teva, which had forty products approved before Cinqaero® and GlaxoSmithKline, with thirty approvals before the application of Strimvelis®. In total, the sample consist of 98 applicants that have had no (successful) past experiences with the EMA before submitting their application. Noticeable is that companies that almost exclusively focus on producing generics, such as Intas and Teva, can gain experience through past approvals much quicker than companies that traditionally focused on innovative products, such as Eli Lilly. Lilly received its first approval 15 years before Intas, but at the beginning of 2017 Lilly had 21 products approved (on average one per year), while Intas had received marketing authorization for 20, averaging 3.33 products per year.

Graph 3: Development of past approvals for the ten companies with most past approvals



Current applications. In previous research, the current experience of an applicant with the FDA proved to positively influence review times, speeding up review times with six months per additional application (Olson, 1997). Current applications are measured by the number of ongoing applications at the time of the application submission of the new product. In the sample, an applicant had on average 0.78 review procedures ongoing at submission date of new product application. The highest current experience is for Teva pharmaceuticals and Novartis, who both had nine new procedures started during the review process of Nexobrid® between November 2010 and September 2012 and Sebivo® between February 2006 and February 2007. A total of 306 review procedures were started without any product reviews ongoing from the same applicant. Graph 4 provides an example of how the current experience of an applicant can change over time, by showing the number of simultaneously ongoing procedures for the companies Johnson and Johnson and GlaxoSmithKline. Note that in the figure GlaxoSmithKline had nine ongoing procedures for a short period in 2007. However, since they have not submitted a MAA during this period, there maximum score is not nine.

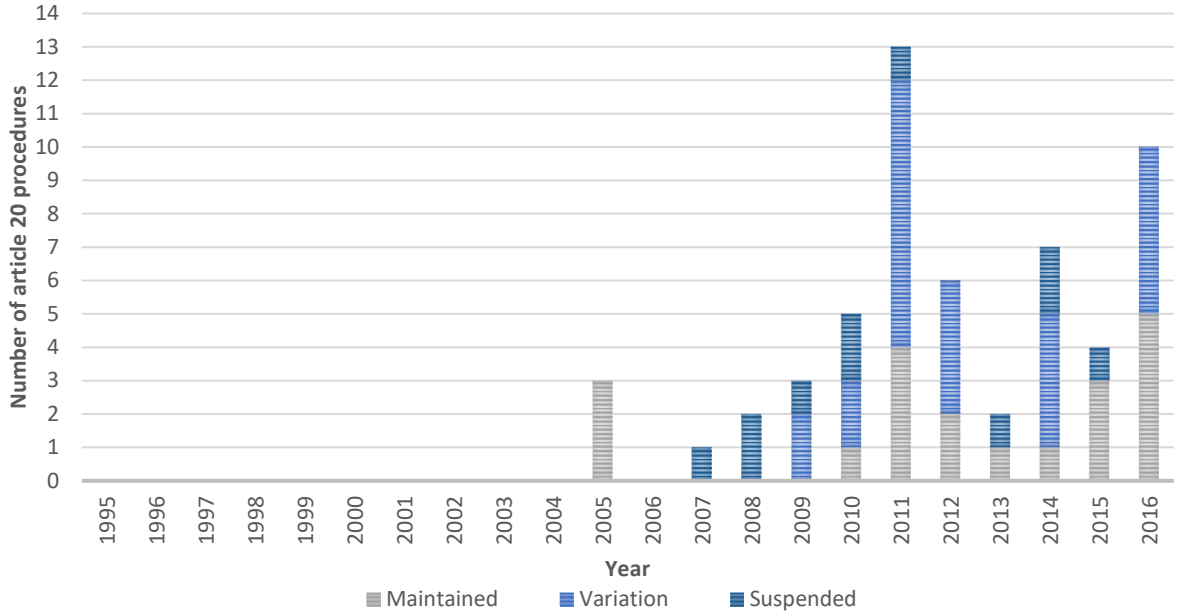
Graph 4: Simultaneously ongoing procedures for Johnson & Johnson and GlaxoSmithKline



4.1.3 Being known for something

Product referrals. Graph 5 visualizes the number of article 20 procedures and their outcomes over time since 2004. A total of 85 centrally authorized products by thirty-two companies have been reassessed under article 20 since 2004. In total, 68 MAAs received approval while the submitting company had been subject to a product referral prior to the opinion date. When looking at absolute numbers, applicants Roche and Novartis both had been subject to five product referrals, but since they have both had a high number of previous approvals, there proportion of referrals was only 29.4% and 10.6% respectively. Bioton, Celgene and TMC pharma held the highest proportion of referrals, having had there only centrally authorized product at that time reassessed on its risk-benefit ratio.

Graph 5: Yearly number of article 20 procedures and outcome



Safety withdrawal. Eleven safety withdrawals have been included to compose the variable if an applicant was involved in a product safety withdrawal before submission of the MAA. These eleven withdrawals have been visualized in figure 6. Eight different companies were responsible for these withdrawals, mostly companies with many applications such as Pfizer (trovafloxacin, valdecoxib, sitaxentan sodium), Merck (laropiprant) and Sanofi Aventis (rimonabant). This resulted in a total of twenty-nine MAAs submitted by an applicant which had previously been involved in a safety withdrawal.

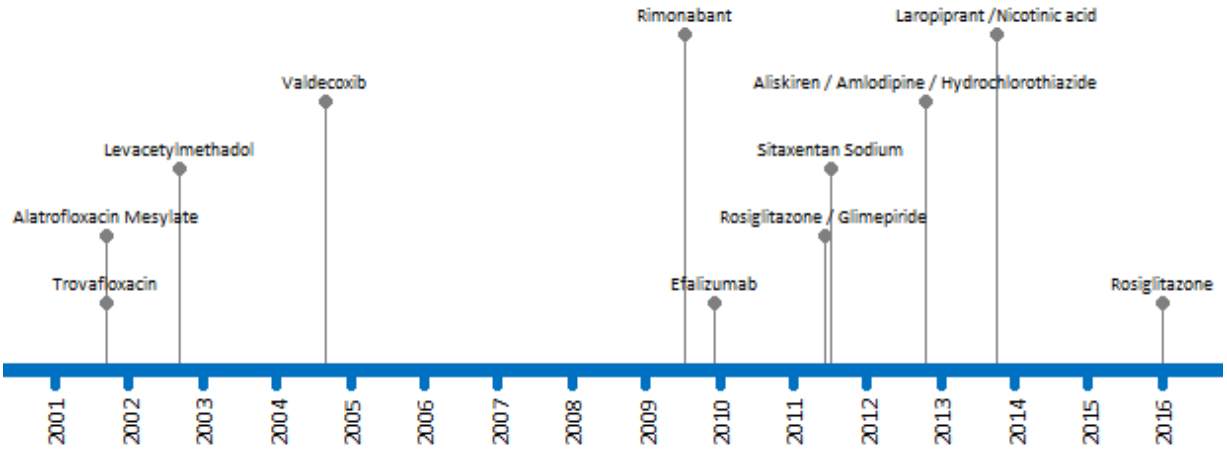
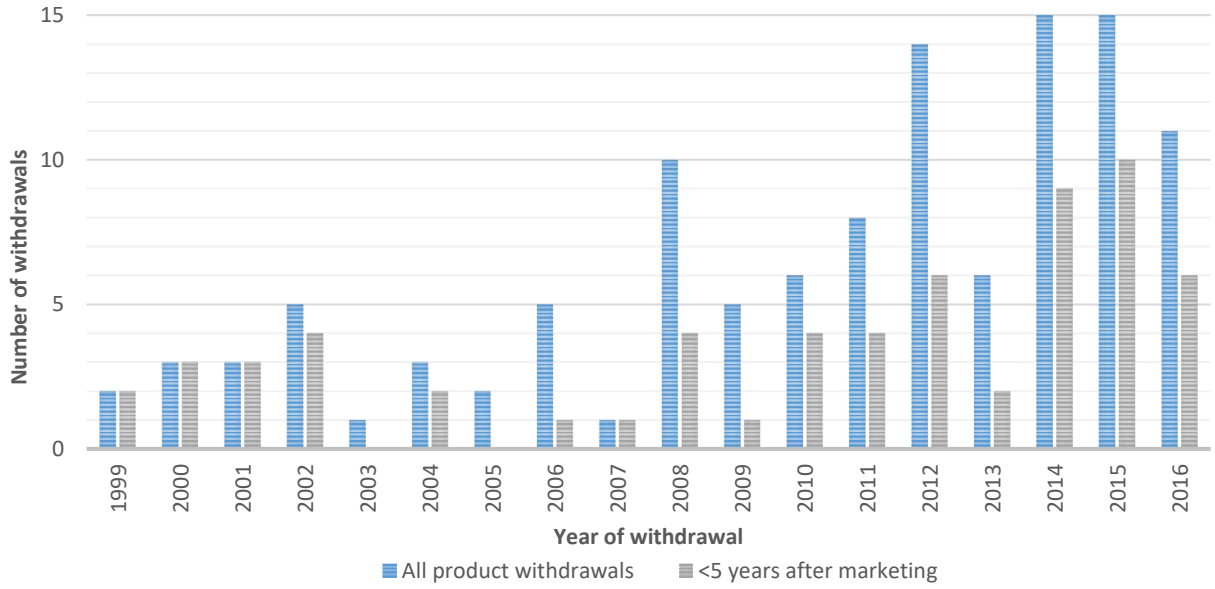


Figure 6: Withdrawals of centrally authorized INNs due to safety issues

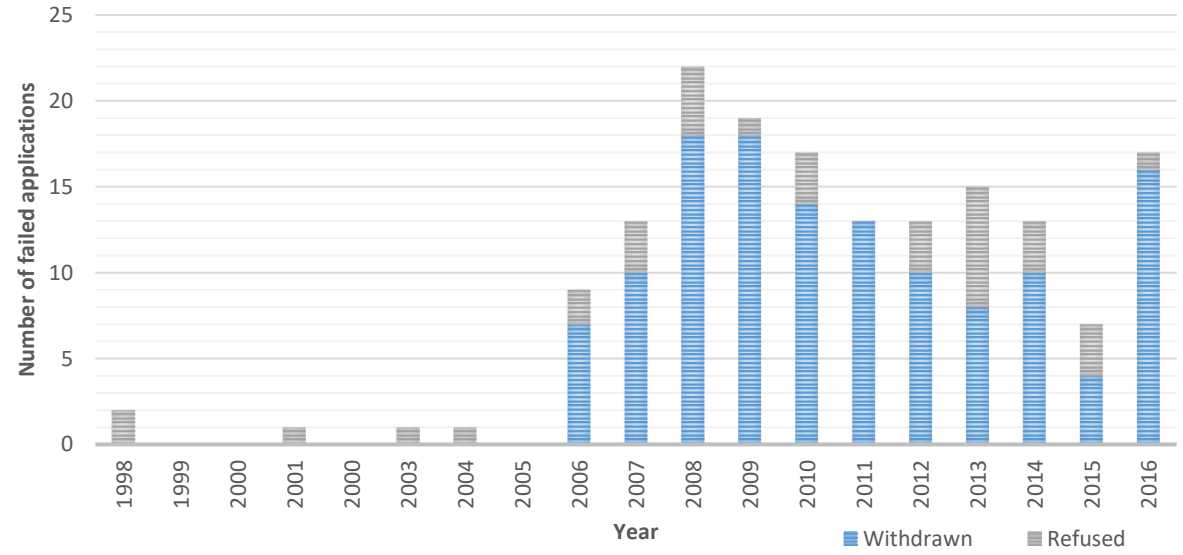
Voluntary withdrawals. The highest absolute number of fast withdrawals is for Teva Pharmaceuticals in 2016, having voluntarily withdrawn six products within five years after initial marketing. This seems a lot, but because Teva places so many products on the market, in fact it withdraws one in every seven products (14.6%). Hence, the proportion of products they withdraw is relatively small compared to other applicants such as Astellas, Celgene, Chiron or Shire who each withdrew on average one in every three products at some point in time. The annual number of total and voluntary withdrawals within five years after initial marketing is shown in graph 6 below.

Graph 6: The number of voluntary product withdrawals per year.



Failed applications. In the sample, there are 360 applications of which the applicant had no failed applications in the past. When looking at the absolute number of failures, it is large companies with pharma sales of over forty billion dollars such as Novartis (10), Sanofi Aventis (4) or Merck and Co. (4) that have frequently failed in the past. Looking at relative amount of failures there is a different image, still with some larger companies, but also including many smaller companies with few or no sales and higher failure rates such as Pharmion (50%), Pharming (50%), Vanda Pharmaceuticals (50%), Celgene (40%) and Bioton (33%). Graph 7 shows the annual number of failed applications split by refusals and withdrawals that have been included in this analysis. As the graph also shows, there is no data available on withdrawals before 2006 and only limited data on refusals before 2006.

Graph 7: The number of refused and withdrawn applications per year



4.1.4 Generalized favorability

***Product suspensions*.** From the 85 reassessment procedures given in graph 5, eleven were suspended. For these products their INN and the referral opinion date is given in figure 7 below.

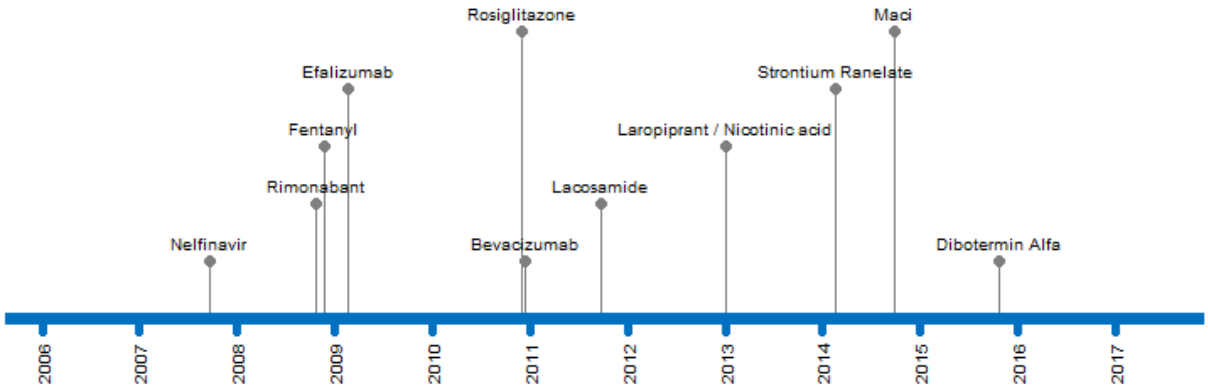
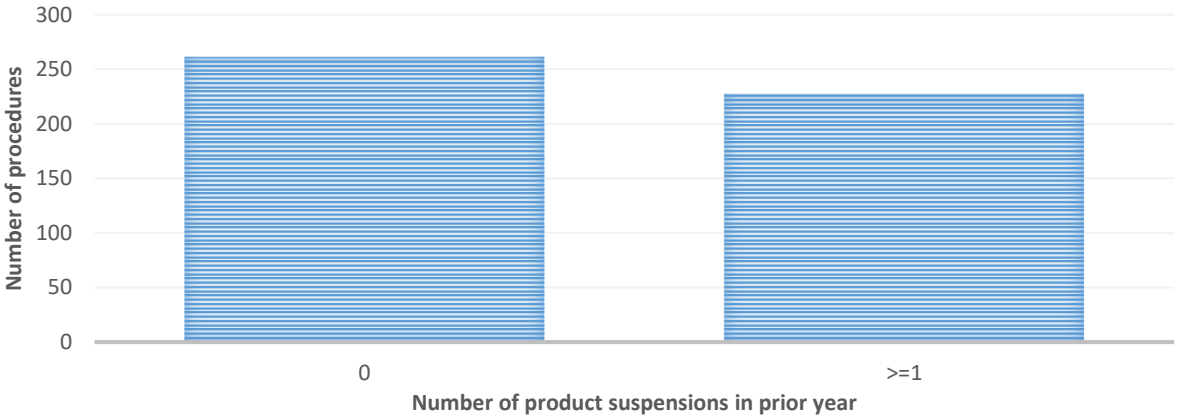


Figure 7: Product suspensions after article 20 reassessment procedure

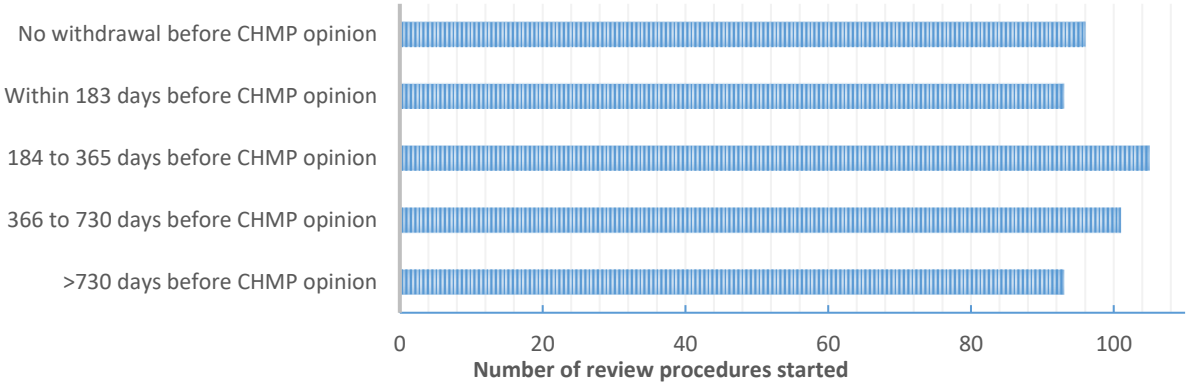
In total, 261 review procedures have not had any product suspension in the year before the application date, as visualized by graph 8. 154 products had one product suspension in the year prior to the application submission. Another 52 had two product suspensions within the year prior to the application submission. Lastly, twenty-one products had three suspended products in the year prior to their submission date. This constitutes a total of 227 procedures that had one or more product suspensions in the year prior to MAA submission.

Graph 8: Number of product suspension in 365 days prior to the submission date of the MAA



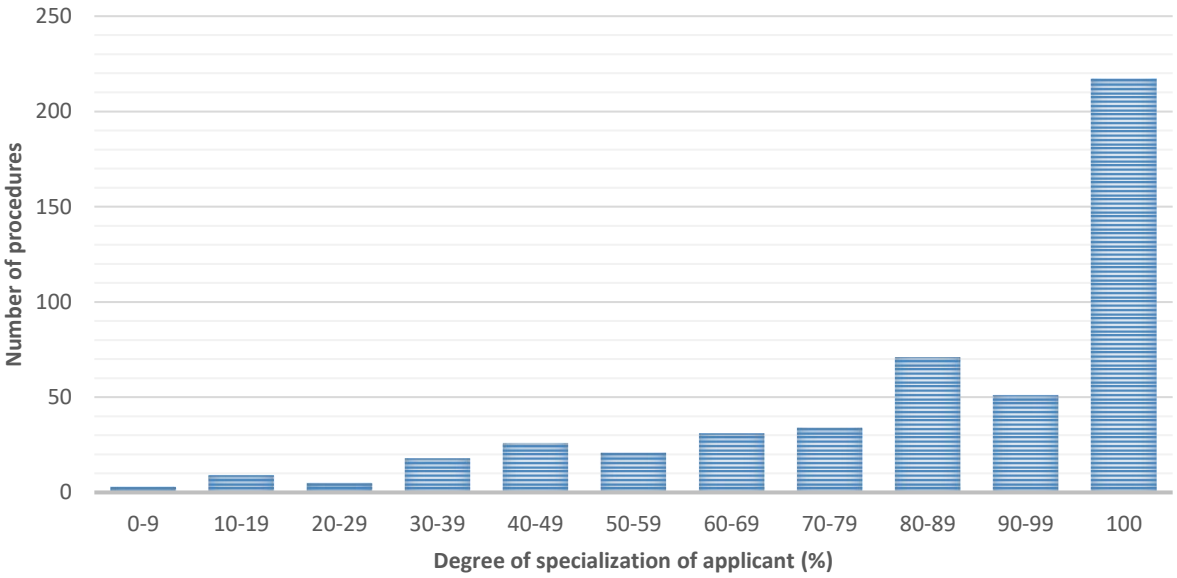
Time since last safety withdrawal. The safety withdrawals included in this research have already been visualized above in figure 6. Since the earliest safety withdrawal received EC decision in March 2001, there are 96 products in the sample that have had no safety withdrawal prior to their submission and received score 1, as visualized in graph 9 below. 93 CHMP opinions followed within 183 days after a central safety withdrawal, 105 between 184 and 365 days and 101 within one to two year after the latest safety withdrawal. Another 93 products received CHMP opinion without any safety withdrawal in the two years prior. The longest time between a safety withdrawal and application submission was 1774 days for Prolia, which was submitted almost five years after the withdrawal of Valdecoxib.

Graph 9: Distribution of categories of prior safety withdrawal variable



Specialization. The distribution of the degree of specialization amongst applicants for a MAA has been visualized in graph 10. As can be seen, the majority of MAA are submitted by specialized pharmaceutical companies with all off their sales from selling pharmaceutical products (217). Another 122 procedures had been submitted by applicants with between 80% and 99% of their sales from pharmaceutical products. Only 61 procedures were submitted by an applicant that had than 50% of their sales from pharmaceutical products.

Graph 10: Distribution of the degree of specialization of applicants



4.2 Descriptive statistics and correlations

Table 8 below summarizes the number of observations, the mean, standard deviation and minimum and maximum values of all variables were used in the regression analyses.

Table 8: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Median	Max
(DV) Active review time	488	195.95	22.41	69	204	245
(DV) Total review time (only used for robustness check)	488	359.04	115.99	69	344.5	946
(U) Product complexity	488	1,782.71	1,009.39	281	1,622	8,747
(U) Evidence base	488	2,182.64	2,783.59	0	1,354	21,929
(U) Time since FDA approval	488	1.86	1.07	1	1	4
(BK) Past approvals	488	5.49	8.16	0	2	49
(BK) Current applications	488	0.78	1.42	0	0	9
(BKS) Proportion of referrals	488	0.03	0.11	0.00	0.00	1.00
(BKS) Involved in safety withdrawal	488	0.06	0.24	0.00	0.00	1.00
(BKS) Proportion of voluntarily product withdrawals	488	0.04	0.10	0.00	0.00	1.00
(BKS) Proportion of failed applications	488	0.05	0.10	0.00	0.00	0.50
(GF) Product suspension prior year	488	0.47	0.50	0	0	1
(GF) Time since last safety withdrawal	488	3.00	1.40	1	3	5
(GF) Degree of specialization	488	0.83	0.23	0.00	0.90	1.00
Status	488	54.32	26.16	0.00	54.40	202.18
Age	488	69.11	50.84	1	67.5	232
Sales	488	11,618.20	13,217.95	0	6,020.4	58,523
Applied for accelerated assessment	488	0.09	0.29	0	0	1
Origin	488	0.46	0.50	0	0	1

Note: DV = Dependent variable, U = Variables measuring product uncertainty, BK = variables measuring "being known"; BKS = variables measuring "being known for something"; GF = variables measuring "generalized favorability"

Table 9 on the next page provides the associations between all variables that have been included in the statistical analyses. Association between several variables were significant, mostly for the variables product complexity, past approvals and sales. The majority of these significant findings are well below the cutoff of 0.35, which are generally considered weak associations (Taylor, 1990). Some significant moderate associations (between 0.36 and 0.67) are for example between the number of past approvals and current applicants, past approvals and sales, sales and age and sales and safety withdrawals. Finding these moderate associations between variables means that they are likely interdependent (Taylor, 1990). For example having a higher number of product approved is associated with higher sales, or having a higher number of current applications is associated with more approved products. To be certain that collinearity is not significantly affecting my outcomes, I performed a Poisson-regression including all predictor variables as given in the correlation table and computed the variance inflation factors (VIFs) (provided in appendix C). The highest score for the VIF was for sales of pharmaceuticals (2.686), followed by the number of previous approvals (2.557), which is well below generally accepted levels of multi-collinearity (O'brien, 2007).

Table 9: Correlations between variables used in the regression analyses

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
[1] (U) Product complexity																
[2] (U) Evidence base	0.13**															
[3] (U) Time since FDA approval	0	-0.01														
[4] (BK) Past approvals	0.36**	0.07	-0.02													
[5] (BK) Current applications	0.14**	0	0.01	0.46**												
[6] (BKS) Product referrals	0.2**	0.01	0.08	0.2**	0.06											
[7] (BKS) Involved in safety withdrawals	0.14**	0.07	0.07	0.28**	0.06	0.38**										
[8] (BKS) Voluntary withdrawals	0.15**	0	0.06	0.11*	0.02	0.52**	0.18**									
[9] (BKS) Failed applications	0.21**	0.05	0.06	0.17**	0.16**	0.28**	0.06	0.13**								
[10] (GF) Product suspension	0.33**	-0.02	-0.06	0.26**	0.06	0.09*	0.06	0.07	0.24**							
[11] (GF) Time since safety withdrawal	0.19**	0	-0.07	0.16**	0.09*	0.02	0.06	0.05	0.09	0.12**						
[12] (GF) Specialization	0	-0.02	-0.03	-0.06	-0.1*	-0.02	0.04	0.06	0.07	0.18**	0.15**					
[13] Status	0.13**	0	0.12**	0.17**	0.14**	0.01	0.04	0.02	0	0.06	-0.02	0.03				
[14] Age	0.11*	0.13**	0.02	0.36**	0.25**	0.17**	0.31**	0.02	0.02	-0.04	-0.05	-0.34**	-0.06			
[15] Sales	0.38**	0.18**	0.05	0.67**	0.46**	0.34**	0.54**	0.17**	0.32**	0.21**	0.14**	-0.04	0.18**	0.45**		
[16] Applied for AA	0.33**	-0.06	-0.03	0.11*	-0.09	0.03	0.01	0.01	0.05	0.15**	0.09	-0.02	0.07	-0.01	0.06	
[17] Origin	-0.15**	0.01	-0.1*	0.13**	0.13**	-0.01	-0.1*	-0.08	0.09	0	-0.02	0	-0.14**	-0.13**	0.09*	-0.14**

*Note: *p<0.05;**p<0.01; U = Variables measuring product uncertainty, BK = variables measuring "being known"; BKS = variables measuring "being known for something"; GF = variables measuring "generalized favorability"*

4.3 Regression models

Firstly, quasi-Poisson regression were performed to compute the influence of the predictors on the CHMP active review time of new pharmaceuticals. To ease interpretation of the effects, all continuous predictor variables with a maximum value of higher than one have been transformed using their natural logarithm. The quasi-Poisson regressions showed that the sample dispersion for all models significantly differed from one ($p<0.001$), which allows for the conclusion that using a quasi-Poisson model over a Poisson model is justified for this dataset. To test the goodness of fit the residual deviance of the models have been compared to the maximum deviance of the model using a Pearson's Chi-squared test as is often done when the independent variable mean is larger than one (Wood, 2002). The outcome of the goodness of fit tests were that all model outcomes were statistically significant ($p<0.001$), which indicates that the models do not fit the data very well. Because the goodness of fit tests showed that the model fit was lacking, a negative binomial regression was performed. The negative binomial regression showed slightly better log-likelihood outcomes and therefore the results of the negative binomial regression are explained in this section. The regression outcomes of the quasi-Poisson models are provided in appendix D.

4.3.1 Negative binomial regressions

Baseline model. Table 10 shows the outcomes of the negative binomial regressions using active review time as dependent variable. The theta parameter is significantly different from one for all models ($p<0.001$), indicating that the negative binomial method is preferred over the quasi-Poisson. The baseline model (controls), only including the predictors that control for product uncertainty and firm characteristics found significant positive effects for the complexity of the product ($\beta=0.040$; $p<0.001$). Theoretically, I argued that complex products are reviewed in the product label in a larger number of words. From the regression with only controls follows that the number of words indeed has a significant positive effect, meaning that more complex products are subject to longer review times. The β-coefficient is rather small, but should be interpreted as any 1% increase in the number of words to describe pharmacological properties and chemical structures results in a 0.04% increase in active review time.

The variable measuring product uncertainty by the time since FDA approval, shows a significant negative effect ($\beta=-0.033$; $p<0.05$). In short, this means that product that have been approved by the FDA in the 183 days before CHMP opinion, have review times that are $e^{-0.033} = 0.968$ times shorter than products that have not received FDA approval at any point before CHMP opinion. These findings correspond with the idea that FDA approval reduces product uncertainty. However, the effect size and significance of the effect decreases when the time since FDA approval increases to more than 183 days. Lastly, applying for accelerated assessment results multiplies active review time by $e^{-0.166} = 0.847$, and as expected shortens active review times with 15.3%.

Being known. Models 1 and 2 both test the effect that being known by the CHMP has on the active review time. Model 1 tests how the number of past approvals effects active review time and the model shows a small but almost significant positive effect ($\beta=0.013$; $p=0.051$). These results are close to being significant at conventional levels. I hypothesized that more previous approvals lead to shorter review times, but contrary to this hypothesis, I find that more past approvals probably extend review times. Model 2 tested for the effect of the number of current applications and found a small but positive significant effect ($\beta=0.022$; $p<0.05$), meaning that a 1% increase in the number of applications leads to an extension of review times with 0.022%, when holding all other predictors constant. Similarly to the findings on past approvals, the direction of the effect is contradictory with the expected direction based on hypothesis 2.

When testing both previous approvals and current applications in one model the effect size of both previous approvals ($\beta=0.010$; $p=0.126$) and current applications ($\beta=0.019$; $p=0.062$) slightly decreases and becomes insignificant. However, when looking at the R^2 and loglikelihood we see an increased

model fit compared to the models testing only control variables and the models testing the effects of being known separately.

Being known for something. Hypothesis 3 and 4 emphasize on the negative impact that being known for quality issues or product discontinuities might have on active review time. Outcomes of the negative binomial models that have tested for product referrals, firm related safety withdrawals, voluntary withdrawals and failed applicants are given in table 10. None of the models reaches conventional levels of statistical significance for the estimated effect sizes of indicators for quality issues of product discontinuities. However, indicators for whether applicants have previously been involved in product safety withdrawals (β=0.039; p=0.077) and the average number of failed applications (β=0.082; p=0.099) are close to conventional levels of significance. All indicators have positive β-coefficients, which indicates that the hypothesized direction of the effect, that being known for quality issues or product discontinuities increases review time, is likely to be correct.

When comparing model fit we can see a small improved in R^2 and loglikelihood scores compared to the model that only includes controls. Including all indicators in the model again estimates similar levels of significance for being known for safety withdrawals (β=0.041; p=0.075) and failed applications (β=0.086; p=0.092). Because three out of four indicators are scored in similar fashion (only model 4 has a binary variable) and the score expresses the proportion of contact with the CHMP that a quality issue or product discontinuity occurred, effect sizes can be compared. It appears that the number of failed applications has the largest impact on CHMP active review time out of the predictors that applicants can be known for.

Generalized favorability. Hypothesis 5 and 6 predicted that the generalized favorability is negatively affected by general safety issues and positively influenced by the degree of specialization of the applicant, which has been tested by models 7, 8 and 9. Model 7 tested the impact of the number of product suspensions in the year before CHMP opinion on active review time. Consistent with what was hypothesized, the estimated coefficient is significantly and positive (β=0.079; p<0.001). This implies review times extend after product suspensions. Having one or more product suspensions in the year prior to CHMP opinion extends active review time with on average by 8.2% ($e^{0.079} = 1.082$).

Model 8 tested the impact of safety withdrawals in the period before application submission and found a positive significant effect (p<0.001) for all categories compared to no safety withdrawal in the prior years. Receiving CHMP opinion within 183 days after a safety withdrawal multiplies the expected number of days of active review time with $e^{0.095} = 1.1$. The effect of a recent safety withdrawal seems to decrease over time, as the effect sizes for the categories 183 to 364 days (β =0.086; p<0.001), 365 to 730 days (β=0.092; p<0.001) and >730 days (β=0.081; p<0.001) are lower than the coefficient for first category. Model 9 tested for the effect of the degree of specialization of the applicant, which showed to be positive and significant (β=0.051; p<0.05). This indicates that a 1% increase in the degree of specialization extends the active review time with 0.051% which is contrary to what was hypothesized in hypothesis 6.

Table 9 showed that there is a significant association between product complexity and the number of product suspensions (r=0.26; p<0.01) and safety withdrawals (r=0.16; p<0.01). This might explain why the product complexity becomes insignificant after introducing the variables measuring general safety issues. However, because the size of the association is small, there could also be an interaction between the complexity and safety issues. To test these possible interactions, separate negative binomial regressions were performed with only the predictors product complexity and product referrals, and only product complexity and safety withdrawals. These regressions showed significant negative effects for the interaction between complexity and if a product suspension had occurred in the prior year (β=-0.050; p<0.05) and product complexity and no prior safety withdrawals (β=-0.033; p<0.01). No significant effects were found between complexity and the other categories of safety

withdrawals (p>0.05). A possible explanation could be that after product suspensions and since the first safety withdrawal in 2001, more complex product have been submitted for review. Another explanation is that since the first safety withdrawal, the number of words to describe pharmacological properties in the product label has increased because of more stringent regulations, as emphasized by Maor (2011).

When comparing the model fit of models 7, 8 and 9 by looking at the R^2 and loglikelihood scores, it can be concluded that the predictor 'time since the last safety withdrawal' explains most of the variance in the active review time. When comparing models 1 till 9, it seems that the indicators for the applicant's generalized favorability are best in explaining the review times. Models that combine the indicators measuring the effects of being known and being known for something, both reach an R^2-score of 0.207, compared to a R^2-score of 0.274 for the model testing all predictors of generalized favorability. To test the goodness-of-fit of all models, the residual deviance has been compared to a chi-squared distribution of the degrees of freedom (Cameron & Windmeijer, 1996). None of the goodness-of-fit tests had significant outcomes at the 5% level, with model 8 having the lowest p-value (p=0.052) and model 3 the highest (p=0.080). These outcomes mean that when using a cutoff of 5%, the models provide an acceptable fit with the data.

Table 10: Coefficients, standard errors and model fit of negative binomial regressions

	Dependent variable: Active review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(U) Product complexity (ln)	**0.040*** **	**0.031** **	**0.037*** **	**0.031** **	**0.039*** **	**0.037*** **	**0.038*** **	**0.036*** **	**0.034** **	0.014	0.003	**0.036*** **	-0.011	-0.013
	(0.010)	**(0.011)**	**(0.010)**	**(0.011)**	**(0.011)**	**(0.010)**	**(0.010)**	**(0.011)**	**(0.011)**	(0.011)	(0.012)	**(0.010)**	(0.012)	(0.012)
(U) Evidence base (ln)	-0.002	0.0001	-0.0003	0.001	-0.002	-0.001	-0.002	-0.002	-0.002	0.001	0.002	-0.001	0.004	0.005
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
(U) <183 days since FDA approval	**-0.033** **	**-0.035** **	**-0.035** **	**-0.036** **	**-0.034** **	**-0.034** **	**-0.034** **	**-0.034** **	**-0.034** **	**-0.027***	**-0.029***	**-0.031***	**-0.024***	**-0.026***
	(0.013)	**(0.013)**	**(0.013)**	**(0.013)**	**(0.013)**	**(0.013)**	**(0.013)**	**(0.013)**	**(0.013)**	**(0.012)**	**(0.012)**	**(0.013)**	**(0.012)**	**(0.012)**
(U) 183 to 364 days since FDA approval	0.020	0.019	0.018	0.017	0.019	0.017	0.019	0.018	0.015	0.023	0.025	0.024	0.027+	0.023
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
(U) >365 days since FDA approval	0.0003	0.0004	-0.0001	0.00001	-0.0002	-0.00005	-0.001	-0.001	-0.001	0.008	0.002	0.0003	0.007	0.006
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
(BK) Past approvals (ln)		0.013+		0.010										-0.004
		(0.006)		(0.007)										(0.007)
(BK) Current applications (ln)			**0.022***	0.019+										**0.022***
			(0.010)	(0.010)										**(0.010)**
(BKS) Product referrals					0.021				-0.063					-0.059
					(0.049)				(0.060)					(0.057)
(BKS) Involved in safety withdrawal						0.039+			0.041+					0.032
						(0.022)			(0.023)					(0.023)
(BKS) Voluntary withdrawals							0.056		0.072					0.059
							(0.050)		(0.058)					(0.055)
(BKS) Failed applications								0.082+	0.086+					0.019
								(0.050)	(0.051)					(0.049)
(GF) Product suspension prior year										**0.079*** **			**0.057*** **	**0.057*** **
										(0.016)			**(0.017)**	**(0.017)**
(GF) <183 days since last safety withdrawal											**0.095*** **		**0.074*** **	**0.077*** **
											(0.017)		**(0.018)**	**(0.018)**
(GF) 183 to 364 days since last safety withdrawal											**0.086*** **		**0.070*** **	**0.070*** **
											(0.017)		**(0.017)**	**(0.017)**
(GF) 365 to 730 days since last safety withdrawal											**0.092*** **		**0.080*** **	**0.080*** **
											(0.017)		**(0.017)**	**(0.018)**
(GF) >730 days since last safety withdrawal											**0.081*** **		**0.066*** **	**0.066*** **
											(0.017)		**(0.017)**	**(0.017)**

Table 10 (continued): Coefficients, standard errors and model fit of negative binomial regressions

	Dependent variable: Active review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(GF) Specialization												**0.051***	0.020	0.018
												(0.023)	(0.022)	(0.022)
Status (ln)	-0.004	-0.005	-0.005	-0.005	-0.004	-0.004	-0.003	-0.003	-0.002	-0.004	-0.002	-0.004	-0.002	-0.002
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Age (ln)	-0.002	-0.004	-0.003	-0.005	-0.002	-0.003	-0.001	-0.001	0.0001	-0.001	-0.0003	0.002	0.001	0.002
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Sales (ln)	-0.0004	-0.002	-0.002	-0.003	-0.0005	-0.001	-0.001	-0.001	-0.002	-0.0003	-0.0003	-0.0003	-0.0001	-0.002
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Applied for AA	**-0.168*** **	**-0.169*** **	**-0.163*** **	**-0.164*** **	**-0.168*** **	**-0.168*** **	**-0.167*** **	**-0.168*** **	**-0.169*** **	**-0.175*** **	**-0.168*** **	**-0.166*** **	**-0.171*** **	**-0.166*** **
	(0.019)	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.018)**	**(0.019)**	**(0.018)**	**(0.018)**
Origin	-0.003	-0.005	-0.006	-0.007	-0.003	-0.004	-0.002	-0.005	-0.004	-0.004	-0.002	-0.003	-0.003	-0.004
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Constant	**5.047*** **	**5.107*** **	**5.068*** **	**5.114*** **	**5.052*** **	**5.066*** **	**5.050*** **	**5.065*** **	**5.077*** **	**5.187*** **	**5.199*** **	**5.010*** **	**5.257*** **	**5.272*** **
	(0.079)	**(0.084)**	**(0.079)**	**(0.084)**	**(0.080)**	**(0.079)**	**(0.079)**	**(0.079)**	**(0.080)**	**(0.082)**	**(0.080)**	**(0.080)**	**(0.085)**	**(0.087)**
Observations	488	488	488	488	488	488	488	488	488	488	488	488	488	488
R^2	0.199	0.199	0.210	0.210	0.201	0.199	0.198	0.198	0.198	0.205	0.245	0.199	0.245	0.261
Log Likelihood	-2,204.3	-2,202.3	-2,201.8	-2,200.6	-2,204.2	-2,202.7	-2,203.6	-2,202.9	-2,200.6	-2,192.9	-2,185.7	-2,201.9	-2,179.3	-2,175.1
Theta	**144.3*** **	**146.5*** **	**147.1*** **	**148.5*** **	**144.4*** **	**146.0*** **	**145.0*** **	**145.8*** **	**148.4*** **	**158.4*** **	**168.0*** **	**147.0*** **	**177.8*** **	**184.4*** **
	(17.031)	**(17.424)**	**(17.527)**	**(17.776)**	**(17.050)**	**(17.329)**	**(17.158)**	**(17.294)**	**(17.746)**	**(19.586)**	**(21.381)**	**(17.515)**	**(23.323)**	**(24.634)**
AIC.	4,430.65	4,428.786	4,427.659	4,427.314	4,432.415	4,429.478	4,431.350	4,429.898	4,431.383	4,409.819	4,401.504	4,427.891	4,392.751	4,396.229

*Note: +p<0.1; *p<0.05; **p<0.01; ***p<0.001; (U) = variables measuring product uncertainty, (BK) = variables measuring being known, (BKS) = variables measuring being known for something; (GF) = variables measuring generalized favorability*

4.3.2 Proportional hazard regression

Baseline model. Table 11 gives the regression coefficients, standard errors and model fit for cox models estimating effect sizes based on all positive CHMP opinions. The coefficients of cox models can best be interpreted by taking its exponent, which gives the average hazard ratio of the corresponding predictor variable. This means negative coefficient scores results in $e^{<0} < 1$, or in other words a decreased approval hazard (or an increased hazard of remaining under review). This equals an increased active review time and vice versa for positive coefficient scores.

The baseline model with only control variables shows a significant effect for having applied for accelerated assessment (p<0.001). A product which has been submitted for accelerated assessment has an average hazard ratio of approval of $e^{0.892} = 2.435$. This means that compared to the baseline hazard, having applied for accelerated approval increases the hazard of approval with 143.5%. However, as mentioned before, this effect is the average of the effect size over time. Hence, the hazard is likely to be higher at day 0 and lower after day 180.

Being known. Model 1 shows the predicted effect of past approvals, which has a very small insignificant effect size ($e^{-0.007} = 0.993$; p=0.332). Similar to the negative binomial regression, even despite the fact that the coefficient is not significant at the 5% level, the direction of the effect is contrary to what was hypothesized in hypothesis 1. The estimated effect is that more previous approvals slightly decrease the hazard of approval. Model 2 estimated the effect of current approvals, which also showed to have a small insignificant effect ($e^{-0.033} = 0.968$; p=0.362). Like the effect of past approvals, the estimated effect of current applications was in the opposite direction as hypothesized. According to the proportional hazard methods, having experience with the regulators does not significantly speed up the active review time of the CHMP.

When looking at the model fit of model 1 and 2 we can see that both R^2 and loglikelihood score barely improve after introducing the predictors for being known. To test the hypothesis that all variables in the model have estimated β-coefficients that are significantly different from zero, Likelihood Ratio, Wald and score tests were executed. All three tests had significant outcomes (p<0.001) and so the omnibus null hypothesis could be rejected.

Being known for something. The models 3 till 6 estimated the effects of applicants that are known for product discontinuities and quality issues. No significant results were obtained at the 5% level, but the estimated effects of having been involved in product safety withdrawals ($e^{-0.295} = 0.745$; p=0.148) and the number of failed applications ($e^{-0.864} = 0.421$; p=0.063) came close to significant levels. Despite the insignificance of the predicted effects, it indicates that applicants that have been involved in a safety withdrawal have an increased hazard of remaining under review of 25.5%. The estimated direction of the effect is negative for all four predictors that indicate if applicants are known by the CHMP for their past performances. This corresponds with hypothesis 3 and 4, that emphasized that applicants that are known for quality issues or product discontinuities are subject to longer review times. Because the number of failed applications is measured by a percentage, every 1% increase in failed applications increases the hazard of remaining under review by 0.421%, indicating that an applicant with 50% of its applications failed has an increased hazard of remaining under review of 21.05% compared to an applicant with no failed applications.

When comparing model fit we see that both the proportion of product referrals and the number of voluntary withdrawals barely explain more of the variance than the baseline model with only control variables. Models 4 and 6 have a slightly better fit to the data, but only improve R^2 with respectively 0.4% and 0.7% compared to the baseline model. Again, Likelihood Ratio, Wald and score tests were executed and all had significant outcomes (p<0.001).

Generalized favorability. Model 7 tested the effects of product suspensions on active review time and shows that a product suspension after a risk-benefit reassessment in the year prior to the review, significantly decreased the hazard of receiving approval by (β=-0.431; p<0.001). This translates to an increase in the hazard of remaining under review of 35% ($e^{-0.431} = 0.650$) for applications that received CHMP opinion within one year after the latest safety withdrawal.

The effect of the time since the last safety withdrawal has been tested in model 8 and shows showed that products that underwent review within 183 days after a safety withdrawal had a significantly decreased hazard ratio (β=-0.516 ;p<0.01). Meaning that, holding all other predictors constant, the hazard of remaining under review is increased with 41.3% ($e^{-0.516} = 0.597$), compared to no prior safety withdrawals. Having had a safety withdrawal within six to twelve months before the product review resulted in increased hazard of 35.9% ($e^{-0.445} = 0.641$; p<0.01), compared to no prior safety withdrawal. The negative effect of prior safety withdrawals reduces over time, as third category reported a hazard ratio of $e^{-0.364} = 0.695$ (p<0.05). The category that compared approvals with no safety withdrawal in the prior two years to the baseline category (no prior safety withdrawals) even showed an increased hazard of approval ($e^{0.118} = 1.125$). However, this estimate was not nearly significant at the 5% level (p=0.486) and therefore it is hard to draw any conclusions from this last category. In general, these findings on prior safety withdrawals are in accordance with hypothesis 5, that safety issues negatively affect the active review time.

Model 9 tested the effect of the degree of specialization and showed a negative but insignificant effect on CHMP review times (β=-0.266; p=0.206). The model combining models 7, 8 and 9 has an improved model fit compared to all other predictors independently, which indicates that most of the variance in the active review time is explained by the predictors measuring general safety issues. Lastly, Likelihood Ratio, Wald and score tests were performed and showed to be significant (p<0.001). However, as already pointed out in the methods section, the proportional hazards assumption estimates the predictors in the regression assuming that the hazards remain constant over time. Since this assumption is violated for several variables, the effect sizes that are found for these variables are an average of the effect size over time. This means that variables that have a changing β-coefficient over time are estimated with the average β-coefficient, which might cancel out the effect of the predictor, showing lower significance levels than might be the case when introducing a cutoff point.

Table 11: Coefficients, standard errors and model fit of proportional hazards models

	Dependent variable: Active review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(U) Product complexity (ln)	-0.174+	-0.140	-0.164+	-0.142	-0.160	-0.155	-0.162	-0.147	-0.131	0.030	-0.022	-0.162	0.103	0.116
	(0.099)	(0.105)	(0.100)	(0.105)	(0.102)	(0.100)	(0.100)	(0.100)	(0.102)	(0.112)	(0.115)	(0.099)	(0.121)	(0.124)
(U) Evidence base (ln)	-0.028	-0.036	-0.034	-0.038	-0.029	-0.031	-0.028	-0.026	-0.028	-0.047	-0.047	-0.030	-0.060	-0.064
	(0.039)	(0.040)	(0.040)	(0.040)	(0.039)	(0.040)	(0.039)	(0.039)	(0.040)	(0.040)	(0.040)	(0.039)	(0.040)	(0.042)
(U) <183 days since FDA approval	0.110	0.113	0.117	0.117	0.115	0.116	0.112	0.113	0.115	0.081	0.132	0.100	0.112	0.124
	(0.118)	(0.118)	(0.119)	(0.119)	(0.118)	(0.118)	(0.118)	(0.118)	(0.118)	(0.119)	(0.119)	(0.118)	(0.120)	(0.120)
(U) 183 to 364 days since FDA approval	-0.163	-0.159	-0.157	-0.156	-0.152	-0.125	-0.152	-0.142	-0.105	-0.180	-0.082	-0.178	-0.105	-0.043
	(0.153)	(0.153)	(0.153)	(0.153)	(0.154)	(0.155)	(0.154)	(0.153)	(0.155)	(0.153)	(0.155)	(0.153)	(0.156)	(0.158)
(U) >365 days since FDA approval	-0.073	-0.080	-0.076	-0.080	-0.064	-0.067	-0.060	-0.069	-0.058	-0.147	-0.089	-0.084	-0.154	-0.132
	(0.143)	(0.143)	(0.143)	(0.143)	(0.144)	(0.143)	(0.144)	(0.143)	(0.144)	(0.145)	(0.144)	(0.144)	(0.146)	(0.147)
(BK) Past approvals		-0.007		-0.005										0.005
		(0.007)		(0.007)										(0.008)
(BK) Current applications			-0.033	-0.023										-0.040
			(0.036)	(0.038)										(0.039)
(BKS) Product referrals					-0.290				0.353					0.606
					(0.459)				(0.595)					(0.606)
(BKS) Involved in safety withdrawal						-0.295			-0.294					-0.380
						(0.204)			(0.210)					(0.238)
(BKS) Voluntary withdrawals							-0.403		-0.456					-0.668
							(0.515)		(0.632)					(0.655)
(BKS) Failed applications								-0.864+	-0.876+					-0.518
								(0.464)	(0.478)					(0.479)
(GF) Product suspension prior year										**-0.431*** **			**-0.367** **	**-0.363** **
										(0.111)			**(0.115)**	**(0.119)**
(GF) <183 days since last safety withdrawal											**-0.516** **		-0.333+	-0.343+
											(0.170)		(0.177)	(0.180)
(GF) 183 to 364 days since last safety withdrawal											**-0.445** **		-0.314+	-0.292+
											(0.162)		(0.164)	(0.165)
(GF) 365 to 730 days since last safety withdrawal											**-0.364***		-0.261	-0.263
											(0.174)		(0.175)	(0.176)
(GF) >730 days since last safety withdrawal											0.118		0.244	0.265
											(0.169)		(0.172)	(0.175)

Table 11 (continued): Coefficients, standard errors and model fit of proportional hazards models

	Dependent variable: Active review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(GF) Specialization												-0.266	-0.145	-0.122
												(0.210)	(0.216)	(0.219)
Status (ln)	0.002	0.005	0.005	0.007	-0.002	0.001	-0.003	-0.005	-0.007	-0.006	-0.009	-0.003	-0.021	-0.027
	(0.068)	(0.069)	(0.069)	(0.069)	(0.069)	(0.069)	(0.069)	(0.069)	(0.069)	(0.069)	(0.069)	(0.069)	(0.070)	(0.071)
Age (ln)	0.018	0.031	0.023	0.031	0.021	0.031	0.014	0.007	0.012	0.012	0.014	-0.004	-0.005	-0.007
	(0.057)	(0.058)	(0.057)	(0.058)	(0.057)	(0.057)	(0.057)	(0.057)	(0.059)	(0.056)	(0.056)	(0.059)	(0.059)	(0.062)
Sales (ln)	-0.013	-0.009	-0.009	-0.006	-0.013	-0.010	-0.011	-0.006	-0.001	-0.016	-0.017	-0.012	-0.017	-0.004
	(0.022)	(0.023)	(0.023)	(0.023)	(0.022)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.024)
Applied for AA	**0.892*** **	**0.885*** **	**0.869*** **	**0.871*** **	**0.885*** **	**0.907*** **	**0.885*** **	**0.911*** **	**0.927*** **	**0.996*** **	**0.887*** **	**0.875*** **	**0.941*** **	**0.973*** **
	(0.163)	**(0.163)**	**(0.165)**	**(0.165)**	**(0.163)**	**(0.163)**	**(0.163)**	**(0.164)**	**(0.165)**	**(0.166)**	**(0.166)**	**(0.163)**	**(0.168)**	**(0.172)**
Origin	-0.016	0.003	-0.004	0.007	-0.014	-0.007	-0.022	0.007	0.008	-0.033	-0.019	-0.026	-0.036	-0.022
	(0.095)	(0.097)	(0.096)	(0.097)	(0.095)	(0.095)	(0.095)	(0.095)	(0.096)	(0.095)	(0.095)	(0.095)	(0.096)	(0.098)
Observations	488	488	488	488	488	488	488	488	488	488	488	488	488	488
R^2	0.067	0.069	0.068	0.069	0.068	0.071	0.068	0.074	0.079	0.095	0.113	0.070	0.132	0.142
Log Likelihood	-2520.042	-2519.556	-2519.614	-2519.366	-2519.830	-2518.925	-2519.715	-2518.168	-2516.898	-2512.456	-2507.759	-2519.266	-2502.376	-2499.556

*Note: +p<0.1; *p<0.05; **p<0.01; ***p<0.001; (U) = variables measuring product uncertainty, (BK) = variables measuring being known, (BKS) = variables measuring being known for something; (GF) = variables measuring generalized favorability*

4.3.3 Proportional hazard regression with cutoff at day 200

Table 12 provides the estimated coefficients derived from the proportional hazards regression using a cutoff of 200 days. By introducing the cutoff at day 200 it could be possible to lose explanatory power, since the maximum possible R^2 likely decreases. However, all models testing the hazard of approval with a cutoff have a maximum possible R^2 of 0.986, compared to 1.000 when not having the cutoff. This indicates that using this cutoff only 2.4% of possible variance explained is lost. Additionally, all goodness-of-fit indicators had significant outcomes for estimated models, indicating good model fits.

Baseline model. The model only including control variables shows that product complexity has a significant negative (β=-0.891; p<0.001) effect on the hazard of remaining under review up to 200 days. A 1% increase in the number of words in the product label results in an increased hazard of remaining under review of 0.9% ($e^{-0.815*\ln(1.01)} = 0.991$). Receiving FDA approval within six months before the CHMP opinion has a significant positive effect (β=0.552; p<0.01) on active review time. Having received FDA approval shortly before the CHMP opinion results in a decreased hazard of remaining under review of 73.7%, compared to not having received FDA approval ($e^{(0.552)} = 1.737$). Comparable to the previous regression, applying for accelerated assessment has a significant (β=1.913; p<0.001) positive effect on the active review time of the product. Applying for accelerated assessment decreases the hazard of remaining under review up until day 200 with 577.3% ($e^{(1.813)} = 6.773$).

Being known. When looking at the estimated effects for being known as given by models 1 and 2, the number of past approvals has a negative effect (β=-0.026; p=0.063) that is close to being significant. The effect of past approvals is very small and any 1% increase in the number of past approvals increased the hazard of remaining under review by 0.003% ($e^{-0.026*\ln(1.01)} = 0.9997$). The effect of the number of current applications is negative and insignificant (β=-0.110; p=0.129). The direction of the effect of the number of past approvals and current applications are both negative, which is in the opposite direction as was hypothesized in hypotheses 1 and 2. But because both predictors are (close to being) significant, it is assumed the predicted direction of the effect is correct.

Being known for something. Models 3, 4, 5 and 6 show negative estimated coefficients and most are close to being significant at the 5% level. Therefore I assume the predicted direction of the effects as hypothesized in hypotheses 3 and 4 to be correct. The number of product referrals is the least close to being significant (β=-1.317; p=0.274). Whether applicants have been involved in a safety withdrawal is almost significant at the 5% level (β=-0.856; p=0.069). This would translate to a 1% increase in the proportion of safety withdrawals results in a 0.8% ($e^{-0.856*\ln(1.01)} = 0.992$) increase in the hazard to remain under review for observations up to 200 days. The average number of withdrawals also is close to significant levels (β=-3.219; p=0.052), indicating that a 1% increase in the proportion of safety withdrawal decreases the hazard of approval with 3.2% ($e^{-3.219*\ln(1.01)} = 0.968$). Lastly, the proportion of failed applications is significant at the 5% level (β=-2.404; p<0.05), which indicates that 1% increase in the proportion of failed applications leads to a 2.4% ($e^{-2.404*\ln(1.01)} = 0.976$) increase in the hazard of remaining under review.

Generalized favorability. All indicators measuring generalized favorability independently have a significant effect on the approval hazard up to 200 days. This indicates that a product suspension in the prior year (β=-1.184; p<0.001), the time since the last safety withdrawal (β=-1.399; p<0.001 to β=-1.041; p<0.001) and the degree of specialization of the applicant (β=-0.850; p<0.001) all significantly increase the hazard of remaining under review for applications up to 200 days. The effect size of safety withdrawals decreases as the time since the last safety withdrawal increases. This indicates that the cautiousness of the CHMP decreases when more time has passed since the last safety withdrawal. Similar to the previous regression methods, most of the variance in active review time is explained by the variables measuring generalized favorability.

Table 12: Coefficients, standard errors and model fit of proportional hazards models with cutoff at 200 days of active review

	Dependent variable: Active review time =<200 days = 1 (n = 174); >200 days = 0 (n = 314)													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(U) Product complexity (ln)	-0.891***	-0.785***	-0.864***	-0.780***	-0.848***	-0.845***	-0.841***	-0.807***	-0.764***	-0.418*	-0.327+	-0.815***	-0.029	-0.015
	(0.164)	(0.173)	(0.165)	(0.173)	(0.168)	(0.166)	(0.166)	(0.167)	(0.169)	(0.186)	(0.187)	(0.165)	(0.198)	(0.199)
(U) Evidence base (ln)	0.043	0.027	0.026	0.017	0.036	0.044	0.036	0.039	0.039	-0.015	-0.025	0.038	-0.059	-0.066
	(0.065)	(0.067)	(0.066)	(0.067)	(0.066)	(0.067)	(0.065)	(0.065)	(0.066)	(0.066)	(0.068)	(0.066)	(0.069)	(0.070)
(U) <183 days since FDA approval	0.552**	0.584**	0.578**	0.601**	0.562**	0.559**	0.572**	0.533**	0.545**	0.441*	0.514**	0.534**	0.429*	0.421*
	(0.189)	(0.189)	(0.189)	(0.189)	(0.188)	(0.188)	(0.189)	(0.189)	(0.189)	(0.190)	(0.190)	(0.188)	(0.192)	(0.195)
(U) 183 to 364 days since FDA approval	-0.211	-0.187	-0.196	-0.177	-0.188	-0.152	-0.165	-0.172	-0.108	-0.220	-0.269	-0.297	-0.306	-0.201
	(0.287)	(0.287)	(0.287)	(0.287)	(0.287)	(0.287)	(0.287)	(0.287)	(0.287)	(0.287)	(0.290)	(0.289)	(0.292)	(0.294)
(U) >365 days since FDA approval	0.036	0.031	0.026	0.026	0.053	0.038	0.043	0.080	0.081	-0.079	0.038	0.027	-0.068	-0.062
	(0.252)	(0.252)	(0.252)	(0.252)	(0.253)	(0.252)	(0.252)	(0.252)	(0.252)	(0.253)	(0.253)	(0.253)	(0.254)	(0.255)
(BK) Past approvals		-0.026+		-0.022										0.022
		(0.014)		(0.014)										(0.016)
(BK) Current applications			-0.110	-0.087										-0.151+
			(0.073)	(0.075)										(0.079)
(BKS) Product referrals					-1.317				1.198					1.431
					(1.204)				(1.413)					(1.376)
(BKS) Involved in safety withdrawal						-0.856+			-0.774					-1.002+
						(0.470)			(0.499)					(0.562)
(BKS) Voluntary withdrawals							-3.219+		-3.024+					-2.055
							(1.653)		(1.773)					(1.607)
(BKS) Failed applications								-2.404*	-2.102*					-0.853
								(1.066)	(1.035)					(1.004)
(GF) Product suspension prior year										-1.184***			-1.050***	-1.061***
										(0.199)			(0.221)	(0.232)
(GF) <183 days since last safety withdrawal											-1.399***		-0.835**	-0.882**
											(0.283)		(0.295)	(0.302)
(GF) 183 to 364 days since last safety withdrawal											-1.286***		-0.925***	-0.922***
											(0.266)		(0.265)	(0.265)
(GF) 365 to 730 days since last safety withdrawal											-1.159***		-0.969***	-0.955***
											(0.264)		(0.259)	(0.263)

Table 12 (continued): Coefficients, standard errors and model fit of proportional hazards models with cutoff at 200 days of active review

	Dependent variable: Active review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(GF) >730 days since last safety withdrawal											**-1.041*** **		**-0.780** **	**-0.699** **
											(0.249)		**(0.251)**	**(0.253)**
(GF) Specialization												**-0.850** **	-0.369	-0.355
												(0.315)	(0.337)	(0.341)
Status (ln)	0.018	0.026	0.031	0.036	0.005	0.014	-0.002	0.003	-0.009	-0.002	-0.012	0.016	-0.030	-0.044
	(0.115)	(0.118)	(0.117)	(0.119)	(0.116)	(0.116)	(0.117)	(0.115)	(0.118)	(0.115)	(0.113)	(0.116)	(0.115)	(0.117)
Age (ln)	0.175+	**0.204***	0.185+	**0.208***	0.185+	0.195+	0.151	0.153	0.142	0.112	0.116	0.099	0.057	0.031
	(0.102)	**(0.102)**	(0.102)	**(0.101)**	(0.103)	(0.103)	(0.104)	(0.103)	(0.106)	(0.100)	(0.099)	(0.105)	(0.100)	(0.105)
Sales (ln)	-0.032	-0.015	-0.018	-0.006	-0.029	-0.027	-0.014	-0.017	-0.001	-0.018	-0.025	-0.033	-0.018	0.006
	(0.037)	(0.038)	(0.038)	(0.039)	(0.037)	(0.037)	(0.038)	(0.038)	(0.039)	(0.038)	(0.037)	(0.037)	(0.038)	(0.040)
Applied for AA	**1.913*** **	**1.923*** **	**1.857*** **	**1.885*** **	**1.910*** **	**1.954*** **	**1.887*** **	**1.952*** **	**1.959*** **	**2.126*** **	**2.033*** **	**1.862*** **	**2.180*** **	**2.185*** **
	(0.218)	**(0.219)**	**(0.220)**	**(0.221)**	**(0.219)**	**(0.219)**	**(0.219)**	**(0.221)**	**(0.222)**	**(0.224)**	**(0.226)**	**(0.217)**	**(0.233)**	**(0.237)**
Origin	0.152	0.215	0.193	0.239	0.157	0.164	0.118	0.187	0.161	0.165	0.144	0.140	0.142	0.129
	(0.157)	(0.160)	(0.158)	(0.161)	(0.157)	(0.157)	(0.157)	(0.157)	(0.158)	(0.158)	(0.158)	(0.157)	(0.159)	(0.165)
Observations	488	488	488	488	488	488	488	488	488	488	488	488	488	488
R^2	0.171	0.178	0.175	0.180	0.173	0.178	0.180	0.181	0.193	0.234	0.228	0.182	0.268	0.282
Log Likelihood	-996.084	-994.075	-994.801	-993.341	-995.323	-993.990	-993.447	-992.943	-989.410	-976.618	-978.731	-992.751	-965.701	-961.054

*Note: +p<0.1; *p<0.05; **p<0.01; ***p<0.001; (U) = variables measuring product uncertainty, (BK) = variables measuring being known, (BKS) = variables measuring being known for something; (GF) = variables measuring generalized favorability*

4.3.4 Robustness check: Total review time as dependent variable

Negative binomial regression. Table 13 provides the outcome of the negative binomial regression using the total review time as dependent variable in the regression. Because the sample dispersion again was significantly greater than one (dispersion=27.200; p<0.001) and the theta parameter (13.713 to 15.230; p<0.001) also was significantly distinct from one for all models, executing a negative binomial regression is justified. Regarding the control variables, the following findings were significant for all models. Whether a product has been approved by the FDA within 183 days before CHMP has a significant negative effect on total review time (β=-0.111 to β=-0.100; p<0.001). This indicates that product uncertainty is lower shortly after FDA approval. Applications that received CHMP opinion more than a year after FDA approval, showed a significantly positive effect on total review time (β=0.079 to β=0.090; p<0.05). This indicates that products that have been approved by the FDA for more than a year, have longer total review times than products that have had no FDA approval. The amount of pharmaceutical sales of the applicant had a negative significant effect in all models (β=-0.021 to β=-0.026; p<0.001), which can be explained by the capabilities and network of larger companies to adhere more quickly to the LoQ of the CHMP, as already emphasized on in the method section. Lastly, whether an applicant had applied for AA also showed significant negative effect on total review time (β=-0.470 to β=-0.493; p<0.001).

The variables emphasizing on the effects of being known, being known for something and generalized favorability as the hypothesized reputational judgement showed similar outcomes as with active review time as dependent variable. The current number of applications was significant and positive (β=0.073 to β=0.079; p<0.01) for all models in which it was included. This indicates that the more ongoing applications an applicant has, the longer it takes them to finish the total review procedure. This might be due to the applicant dividing its efforts on multiple applications and therefore be slower in gathering additional clinical data to answer to the LoQ. Being known for quality issues or product discontinuities showed no significant impact on total review time. Only the proportion of failed applications showed levels that approached significance (β=0.174; p=0.164) which increased (β=0.216; p=0.093) when included in the regression simultaneously with other predictors mimicking applicants prior performances.

The generalized favorability of applicants measured by the number of general safety events via product suspensions in the prior year (β=0.112;p<0.001) and time since the last safety withdrawal (β=0.236 to 0.153; p<0.001) showed to significantly slow down total review times. This indicates that the CHMP is more stringent towards applicants when a general safety issue has occurred. The effect size of safety withdrawals decreased over time, which indicates that the more time has passed since a safety withdrawal, the less cautious the regulators have become, or the better adapted the applicants are to new regulations.

Proportional hazard regressions. Appendix E provides the estimated effects of predictors on the total review time using the proportional hazards method. The results are highly similar, but some outcomes were different from the negative binomial regressions. Most noticeable differences were that the effects of past approvals became significant (β=0.018; p<0.05) in the model with all predictors. This indicates that having more past approvals speeds up total review time. Additionally, the effect of whether applicants were previously involved in safety withdrawals became significant (β=-0.454; p<0.05). This could be explained as follows, applicants that were involved in a safety withdrawal might receive a more extensive LoQ by the CHMP to ensure that safety is sufficiently proven at the time the CHMP needs to give an opinion.

Table 13: Coefficients, standard errors and model fit of binomial regressions using total review time as dependent variable

	Dependent variable: Total review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(U) Product complexity (ln)	**0.053***	0.045	0.043+	0.042	**0.056***	0.050+	0.051+	0.046+	0.047+	0.016	-0.029	0.049+	-0.039	-0.040
	(0.026)	(0.028)	(0.026)	(0.028)	**(0.027)**	(0.026)	(0.026)	(0.027)	(0.027)	(0.029)	(0.029)	(0.026)	(0.031)	(0.031)
(U) Evidence base (ln)	-0.010	-0.008	-0.004	-0.004	-0.010	-0.010	-0.010	-0.009	-0.010	-0.005	-0.002	-0.009	-0.0003	0.002
	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)
(U) <183 days since FDA approval	**-0.111*** **	**-0.113*** **	**-0.116*** **	**-0.116*** **	**-0.109*** **	**-0.111*** **	**-0.111*** **	**-0.112*** **	**-0.110*** **	**-0.104****	**-0.102*** **	**-0.109*** **	**-0.100****	**-0.099****
	(0.032)	**(0.032)**	**(0.032)**	**(0.032)**	**(0.032)**	**(0.032)**	**(0.032)**	**(0.032)**	**(0.032)**	**(0.032)**	**(0.031)**	**(0.032)**	**(0.031)**	**(0.031)**
(U) 183 to 364 days since FDA approval	0.033	0.032	0.027	0.026	0.036	0.031	0.032	0.029	0.029	0.037	0.041	0.037	0.041	0.035
	(0.041)	(0.041)	(0.041)	(0.041)	(0.042)	(0.042)	(0.041)	(0.041)	(0.042)	(0.041)	(0.040)	(0.042)	(0.040)	(0.040)
(U) >365 days since FDA approval	**0.081***	**0.081***	**0.081***	**0.081***	**0.083***	**0.081***	**0.080***	**0.079***	**0.080***	**0.093***	**0.087***	**0.081***	**0.092***	**0.092***
	(0.039)	**(0.039)**	**(0.038)**	**(0.038)**	**(0.039)**	**(0.039)**	**(0.039)**	**(0.039)**	**(0.039)**	**(0.039)**	**(0.038)**	**(0.039)**	**(0.038)**	**(0.037)**
(BK) Past approvals (ln)		0.012		0.001										-0.018
		(0.016)		(0.016)										(0.017)
(BK) Current applications (ln)			**0.074****	**0.073****										**0.080****
			(0.025)	**(0.026)**										**(0.025)**
(BKS) Product referrals					-0.085				-0.248					-0.250+
					(0.123)				(0.153)					(0.147)
(BKS) Involved in safety withdrawal						0.032			0.050					0.054
						(0.055)			(0.058)					(0.058)
(BKS) Voluntary withdrawals							0.060		0.161					0.118
							(0.127)		(0.147)					(0.142)
(BKS) Failed applications								0.174	0.216+					0.075
								(0.125)	(0.129)					(0.126)
(GF) Product suspension prior year										**0.112****			0.044	0.050
										(0.042)			(0.043)	(0.044)
(GF) <183 days since last safety withdrawal											**0.236*** **		**0.222*** **	**0.233*** **
											(0.044)		**(0.047)**	**(0.047)**
(GF) 183 to 364 days since last safety withdrawal											**0.230*** **		**0.220*** **	**0.221*** **
											(0.042)		**(0.044)**	**(0.043)**
(GF) 365 to 730 days since last safety withdrawal											**0.174*** **		**0.168*** **	**0.168*** **
											(0.043)		**(0.044)**	**(0.045)**
(GF) >730 days since last safety withdrawal											**0.153*** **		**0.144*** **	**0.144*** **
											(0.042)		**(0.043)**	**(0.044)**

Table 13 (continued): Coefficients, standard errors and model fit of binomial regressions using total review time as dependent variable

	Dependent variable: Total review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(GF) Specialization												0.053	-0.009	-0.011
												(0.058)	(0.057)	(0.057)
Status (ln)	-0.011	-0.012	-0.014	-0.014	-0.012	-0.011	-0.010	-0.009	-0.009	-0.011	-0.011	-0.010	-0.010	-0.014
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.017)	(0.018)	(0.017)	(0.017)
Age (ln)	-0.0004	-0.003	-0.004	-0.004	0.0003	-0.001	0.001	0.002	0.007	0.0005	0.001	0.003	0.001	0.005
	(0.014)	(0.015)	(0.014)	(0.015)	(0.014)	(0.014)	(0.015)	(0.014)	(0.015)	(0.014)	(0.014)	(0.015)	(0.014)	(0.015)
Sales (ln)	**-0.021*** **	**-0.023*** **	**-0.026*** **	**-0.026*** **	**-0.021*** **	**-0.022*** **	**-0.022*** **	**-0.023*** **	**-0.024*** **	**-0.021*** **	**-0.021*** **	**-0.021*** **	**-0.021*** **	**-0.024*** **
	(0.006)	**(0.006)**	**(0.006)**	**(0.006)**	**(0.006)**	**(0.006)**	**(0.006)**	**(0.006)**	**(0.006)**	**(0.006)**	**(0.006)**	**(0.006)**	**(0.006)**	**(0.006)**
Applied for AA	**-0.486*** **	**-0.488*** **	**-0.468*** **	**-0.469*** **	**-0.487*** **	**-0.487*** **	**-0.486*** **	**-0.487*** **	**-0.488*** **	**-0.496*** **	**-0.485*** **	**-0.483*** **	**-0.489*** **	**-0.471*** **
	(0.047)	**(0.047)**	**(0.047)**	**(0.047)**	**(0.047)**	**(0.047)**	**(0.047)**	**(0.047)**	**(0.047)**	**(0.047)**	**(0.045)**	**(0.047)**	**(0.046)**	**(0.045)**
Origin	-0.011	-0.013	-0.020	-0.020	-0.011	-0.012	-0.010	-0.015	-0.012	-0.012	-0.006	-0.011	-0.007	-0.012
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.025)	(0.026)	(0.025)	(0.025)
Constant	**5.822*** **	**5.879*** **	**5.890*** **	**5.895*** **	**5.804*** **	**5.840*** **	**5.824*** **	**5.858*** **	**5.848*** **	**6.012*** **	**6.184*** **	**5.784*** **	**6.248*** **	**6.252*** **
	(0.198)	**(0.212)**	**(0.197)**	**(0.211)**	**(0.199)**	**(0.199)**	**(0.198)**	**(0.199)**	**(0.201)**	**(0.210)**	**(0.201)**	**(0.202)**	**(0.217)**	**(0.223)**
Observations	488	488	488	488	488	488	488	488	488	488	488	488	488	488
R^2	0.199	0.199	0.210	0.210	0.201	0.199	0.198	0.198	0.198	0.206	0.245	0.199	0.246	0.262
Log Likelihood	-2916.387	-2916.112	-2912.009	-2912.006	-2916.153	-2916.221	-2916.281	-2915.435	-2913.978	-2912.872	-2898.638	-2915.963	-2898.277	-2891.368
Theta	**13.713*** **	**13.728*** **	**13.966*** **	**13.966*** **	**13.726*** **	**13.721*** **	**13.719*** **	**13.767*** **	**13.850*** **	**13.920*** **	**14.772*** **	**13.737*** **	**14.805*** **	**15.249*** **
	(0.904)	**(0.905)**	**(0.922)**	**(0.922)**	**(0.905)**	**(0.905)**	**(0.905)**	**(0.908)**	**(0.914)**	**(0.919)**	**(0.978)**	**(0.906)**	**(0.981)**	**(1.012)**
AIC.	5,854.774	5,856.223	5,848.018	5,850.013	5,856.306	5,856.442	5,856.562	5,854.870	5,857.956	5,849.565	5,827.276	5,855.926	5,830.224	5,828.169

*Note: +p<0.1; *p<0.05; **p<0.01; ***p<0.001; (U) = variables measuring product uncertainty, (BK) = variables measuring being known, (BKS) = variables measuring being known for something; (GF) = variables measuring generalized favorability*

4.3.5 Robustness check: Correct for autocorrelation

The outcomes of the regressions using clustered standard errors are given in appendix F. After clustering per applicant to account for correlation that exist between observations of the same applicant, largely the same outcomes are obtained as using the normal negative binomial regression. However, we see a decrease in significance levels for FDA approval within 183 days, but an increase in significance levels for FDA approval within 183 to 365 days. Additionally, the effects of being known are closer to significant levels, which again indicates that a larger number of previous approvals and current applications slows down active review time. Being known for safety withdrawals (β=0.041; $p<0.05$), voluntary withdrawals (β=0.056; $p<0.05$) or having had a high proportion of failed applications (β=0.086; $p<0.05$) all have estimated positive effects and reach significant levels. This indicates that being known for quality issues or product discontinuities does indeed affect the reputational judgement by the CHMP. Regarding the generalized favorability of the applicants based on general safety issues, there are no changes in significance levels after introducing the robust standard errors. The degree of specialization however, becomes insignificant (β=0.051; $p=0.137$). This indicates that the effect generalized favorability is conveyed mostly from general safety issues and more than by the individual level of specialization that the applying company has.

4.3.6 Robustness check: Control for therapeutic area

To test for the effect that the therapeutic area might have on active review times, I firstly performed a univariate regression that estimated the effects of only the therapeutic class on active review times. I found that none of the therapeutic areas significantly affects review times, as shown in table 14 below. However, to comply with Olson (1997), I performed a regression using a dummy variable per therapeutic area. The regression outcomes of the negative binomial regressions using dummies for the therapeutic classes are given in appendix G. None of the therapeutic classes showed significant effect on active review times nor on the significance levels of the indicators of organizational reputation. Additionally, compared to the negative binomial regression without controlling per therapeutic area, I found that the effect of FDA approval within 183 days before CHMP slightly decreases in both effect size and significance levels. The category with FDA approval within 183 to 365 days slows an increase in both effect size and significance. The direction of the effect is positive, which indicates that after controlling for therapeutic area the procedures that received CHMP opinion 183 to 365 days after FDA approval have significantly longer active review time.

Table 14: Univariate regression of therapeutic classes on active review time

Independent variable: Active review time	Coefficient	Standard error
(A) Alimentary tract and metabolism	0.0003	(0.119)
(B) Blood and blood forming organisms	-0.0001	(0.119)
(C) Cardiovascular system	-0.024	(0.120)
(D) Dermatologicals	0.018	(0.124)
(G) Genito urinary system and sex hormones	-0.030	(0.120)
(H) Hormonal preparations	-0.012	(0.123)
(J) Anti-infectives	-0.083	(0.118)
(L) Antineoplastic and immunomodulating agents	-0.033	(0.118)
(M) Musculoskeletal system	0.002	(0.123)
(N) Nervous system	0.002	(0.119)
(P) Antiparasitic products	0.039	(0.166)
(R) Respiratory system	-0.002	(0.122)
(S) Sensory organs	-0.022	(0.121)
(V) Various	-0.033	(0.121)
Constant	5.303***	(0.118)
Observations	488	
Log Likelihood	-2,237.422	
Theta	112.865***	(-11.945)
AIC	4,504.843	

*Note: *p<0.05; **p<0.01; ***p<0.001*

5. Discussion

In this section the scientific contribution and main findings (5.1), theoretical (5.2) and practical implications (5.3) are outlined, a reflection on the research methods and its limitations is provided (5.4) and ideas for future research fields are suggested (5.5).

5.1 Scientific contribution and main findings

5.1.1 Scientific contribution

The research question of this thesis was to what extent different types of organizational reputation influence the length of active review of MAAs by the CHMP. By answering this question, this research provides an empirical contribution to existing knowledge in management literature on organizational judgements and its effects on decision-making under conditions of inherent uncertainty. The contribution of this study lies on clarifying the effects of organizational reputation in social judgements of organizations. In management literature, it has been emphasized that social judgements about organizational legitimacy, reputation and status are being used as heuristics for decision-making (Bitekine, 2011). The decision-makers use these heuristics because they are only boundedly rational and therefore cannot make perfect decisions. While status judgements are primarily used to screen for alternatives, the reputational judgement is especially involved in the decision-making process (Jensen & Roy, 2008). As various scholars have emphasized, reputational judgements provide an indication about expectations and probable outcomes based on previous performances (Lange et al., 2011). Therefore, this research focused on clarifying the effects that organizational reputation has under circumstances of bounded rationality.

To emphasize the effects of organizational reputation, I have chosen to research its effects in the context of pharmaceuticals, were decision-making is always boundedly rational, as product uncertainty can never be completely eliminated before large-scale consuming (Carpenter, 2004). The impact of organizational reputation in the context of regulatory decision-making at the FDA, has been investigated by several scholars (i.e. Olson, 1997; Carpenter et al., 2010). However, none emphasized on the three-dimensionality of reputation as proposed by Lange et al. (2011), who argued that firms can be ‘known’, ‘known for something’ and have a ‘generalized favorability’. Where the majority of the contributions included the familiarity between regulator and applicant (‘known’) and the generalized favorability, only Kim (2012) also included firm-specific prior performances. But as Lange et al. (2011) argued, the evaluation of the familiarity of the applicant is non-judgmental and does not include likelihood estimations about future outcomes. Similarly, generalized favorability is a generic judgement of the regulator and also lacks the specificity to generate expectations of future outcomes for specific applicants. Future expectations are mainly derived from a particular judgement about what the organization is specifically known for (Lange et al., 2011). Hence, the specific traits applicants of MAAs are known for, also provide options for regulators to be distinctive in their judgement.

Prior work focused exclusively on investigating the effects of organizational reputation on decision-making time at the FDA, while scholars indicate that its European counterpart, the EMA, likely also used social judgements (i.e. Hauray, 2017; Tafuri et al., 2014). Investigating the review procedure performed by the CHMP, the decisive body of the EMA, not only allowed for validation of earlier findings, but also allowed to include all three dimensions of organizational reputation in a single and new research setting. Conforming to prior work (see Carpenter et al., 2010; Olson, 1997), I included prior approvals and current applications as indicators of familiarity with the regulator and the process. Measurements of generalized favorability such as the degree of specialization and safety events were replicated from prior work by Olson (1997), Carpenter (2002) and Kim (2012). Since Kim (2012) was the only scholar to include firm-specific indicators, namely involvement in safety withdrawals, I mimicked his approach, but additionally also included three more indicators that emphasized on the prior performance of applicants. Using the three-dimensional approach of reputation allowed for

verification of previous research outcomes and additionally using four indicators to account for what firms are known for, allows for estimating the effects that prior performances have on the organizational judgement that the CHMP conveys when assessing MAAs.

5.1.2 Main findings

Firstly, I found that active review times significantly dependent on product uncertainty, measured by its complexity and whether FDA approval has occurred. Findings that product complexity significantly affect review speed is in line with earlier findings by Hofer et al. (2018), Moffitt (2010) and Olson (1997). It appeared that the more complex a product is, the longer its review generally takes. Receiving FDA approval shortly before the CHMP opinion date, significantly shortens the reviewing process and therefore we can conclude that FDA approval decreases uncertainty in decision-making. However, this effect decreased to insignificant levels when the time between FDA approval and CHMP opinion increases. This could be due to the increased collaboration between the FDA and EMA over the years standardizing clinical trial conducts, scientific advice and strengthening communication (Makuch & Shi, 2013; Tafuri et al., 2014). It is possible that, especially in later years, this increased communication resulted in the CHMP being able to better assess product uncertainty. When the EMA reviews a MAA simultaneously or shortly after the FDA it is possible that the knowledge sharing is better compared to when the period of time between review procedures is more extensive, which could explain the differences in significant effects between FDA approval within 183 days or more.

Focusing on three dimensions of organizational reputation, I established six hypothesis (H1-6) that hypothesized the effects of several aspects of organizational reputation on the social judgement process by the CHMP, measured through the active review time. Hypothesis 1 and 2 focused on the familiarity of the applicant with the regulator and the regulatory process. I was not able to reject nor confirm the first hypothesis due to statistical insignificance of the results, but after introducing clustered standard errors it appears that the number of previous applications likely extends active review times. Hypothesis 2 can be rejected, as the regressions unambiguously showed that the number of current applications extends active and total review times. These findings are contrary to what Olson (1997) and Stern (2017) found, and besides were insignificant in the proportional hazards regressions, which is similar to the findings of Kim (2012).

Hypothesis 3 and 4 focused on whether firms that are specifically known for prior safety issues or product discontinuities are subject to longer review times. Despite lacking clear significant outcomes, I showed that these factors are linked to extension of active review time as they are positively related to the number of days of active review time and were often close to being significant according to conventional statistical levels. These outcomes are an extension on the work by Kim (2012), who found no significant relation between firm-specific safety withdrawals and review times at the FDA. I found that firm-specific performances are noticed by the CHMP, especially prior involvement in safety withdrawals and the proportion of failed applications balance on the edge of statistical significance and therefore are likely to result in longer review times. Finding these results indicates that ‘being known for something’ does affect the reputational judgement by the CHMP as all indicators are unanimously related to longer active review times.

Reputation through generalized favorability was tested by looking at the effects of general safety issues (H5) and the degree of specialization of the applicant (H6). Hypothesis 5 can be confirmed, as I found in all regressions that general safety events extend active review times and are responsible for explaining most of the variance in review times, as corresponds with the findings by Kim (2012). As already explained, it is likely that the regulators behave more cautiously shortly after a general safety issue and takes fewer risks regarding early approval, as emphasized by other scholars (Carpenter, 2002; Carpenter et al., 2010; Maor, 2011). This effect is also displayed by the decreasing effect size as the time since the last safety withdrawal increases. Moreover, it might well be that new guidelines and regulations are installed after each safety withdrawal which requires adaption of both the CHMP and

applicants (Maor, 2011). Hypothesis 6 could not be confirmed as the effect of the degree of specialization had inconclusive findings regarding its significance of the regression outcomes.

5.2 Theoretical implications

During the last decade, the theory of social judgement of organizations has received noticeable attention in several streams of scientific literature. This research proved that social judgements significantly affect decision-making under conditions of inherent uncertainty, as proposed by Bitektine (2011). Where many high-valued theoretical contributions focused on differentiating between legitimacy, status and reputation (e.g. Bitektine, 2011; Rindova et al., 2006), or demarcating reputation and status (e.g. Petkova et al., 2014; Rao, 1994; Rindova et al., 2005), few emphasized on the multidimensionality of organizational reputation as proposed by Lange et al. (2011). This empirical work showed that for the specific context of regulatory decision-making in the European pharmaceutical industry, the ideas of Bitektine (2011) about social judgements and Lange et al. (2011) about the three-dimensionality of organizational reputation, hold. Additionally, I found that any form of reputation whatsoever is likely to extend review time, which indicates that having any reputation at all might be unfavorable for receiving faster approvals, compared to not having reputation. This is contradictory with the ideas of Bitektine (2011), who hypothesized: *"An organization with unknown reputation will be categorized as reputation neutral until more information is available" (p. 165).* Nonetheless, this research focuses specifically on the judgements by a regulator and therefore I cannot say whether this also holds for unregulated markets.

When comparing the three dimensions of organizational reputation, Lange et al. (2011) argued that 'being known' is a nonevaluative judgement. In contrast, I found that 'being known' significantly extends active review time, although the effect sizes were only marginally. However, as judgements in this dimension are theoretically non-evaluative, the significant outcomes indicate that familiarity has an evaluative component. 'Being known for something' is said to be both judgmental and particular for one applicant and therefore likely to constitute the most specific judgment. Although most regressions found no significant results for being known for a specific trait, after introducing clustered standard errors the effects became more clearly visible. This indicates that the organizational judgements are indeed fueled by firm-specific past performances. Lastly, it is clear the effects of the generalized favorability induced by general safety issues have the largest impact on review times . This has theoretically already been discussed by Carpenter (2004) and later proved by Kim (2012), who emphasized on the additional measures by the FDA to protect its own reputation after general safety issues. Although I have not laid emphasis on this, my findings indicate that theoretically similar mechanisms play a role in the European context.

5.3 Practical implications

I added to existing empirical knowledge on regulatory science in the pharmaceutical context gathered by scholars such as Kim (2012), Olson (1997) and Stern (2017). Specifying their contributions according to the dimensions of reputation allowed for combination and replication of their methods and applying these to the, yet under studied, case of MAAs in Europe. Using this approach allows for first cautious conclusions on how the organizational judgement of the CHMP is affected by the reputation of applying companies and how this compares to the FDA. Firstly, I have validated the qualitative work by Hauray (2017), who indicated that the EMA and CHMP are not purely objective in medicine assessment. This work underlines his verdict, showing that the risk-benefit of medicines is partly socio-politically constructed.

Modelling the effects of organizational reputation shows that estimating the effects of social judgements by the CHMP improved model fit, compared to only including control variables that measured product uncertainty and general firm characteristics. This indicate that although only marginally, reputational judgements do play a role in determining active review times. Finding that

several types of reputation significantly extend review times interferes with the 'scientistic view' that medicine regulators adopt in general (Hauray, 2017). This implies that the CHMP should become aware of the effects that socially constructed judgements have on the length of procedures of MAAs. It is undoubtedly clear that product uncertainty, which is constructed by the amount of evidence on product efficacy and safety, influences review times. However, since product uncertainty can never completely be eliminated in drug review, the EMA and CHMP should realize that there will always be some human judgement affecting decision-making.

There can be thought of ways to decrease the effects of social judgement of applicants by the CHMP. Firstly, if the CHMP does not know which applicant has submitted the MAA, there is also no possibility to be influenced by the reputation of that applicant. Most straightforward, this would mean that there should be no contact between (representatives of) the applicant and the CHMP. Practically, this might be rather difficult since the CHMP often files a LoQ, which the applicant is allowed to clarify in a personal meeting. However, as the LoQ is often compiled based on clinical data, it is possible to install an intermediary that censors all links to the applicant in the documents containing the clinical data. In this way the initial safety and efficacy can be judged and assessed independently from the applicant.

For the applicants it is advantageous to know their reputation affects the review time by the CHMP. If an applicant in the past has been involved in many failed applications or safety withdrawals, it could try to submit applications via subsidiaries, or in collaboration with others with a more favorable reputation. Additionally, because applicants know that their general favorability affects the judgement by the CHMP, they could try to emphasize on the organizational differences that exist between them and others in the industry. In this way they could try to avoid a general judgement by the CHMP about their favorability.

Lastly, patients that use products approved by the CHMP should be aware that the EMA is a regulatory agency established to protect patients from possible product hazard (Green & Moore Jr., 1972). As the EMA tries to adequately execute this function the CHMP always faces uncertainty in decision-making and can never completely rule out future safety issues. To easy their decision-making they make use of alternative factors that might indicate the level of product uncertainty, for example based on past performances of applicants. In this way the CHMP tries to minimize the risks associated with new drugs and provide early access to new medicines. But still, patients should understand and accept that drug consuming is never completely without risks, especially shortly after marketing authorization and that past performances of applicants do not provide guarantees for the future.

5.4 Limitations

As any other, this research is subject to certain limitations with regard to research quality and outcomes. This section emphasizes on the choices that have influenced research reliability and validity and how these choices have implications for the research outcomes.

Ensuring reliability was pursued by extensive explanation on data gathering and collecting most data from publicly available sources. However, some data on control variables (e.g. FDA approval dates, applying for AA) was provided by Hoekman, which makes these harder to replicate. Additionally, data on some indicators was only available after a certain year. For example, data on product referrals was only available after 2005 and therefore was assumed that no reputation existed before 2005 on this variable. This might also be the reason why the number of voluntary withdrawals lacked statistical significance on all occasions. Nonetheless, I tried to shape the data in multiple ways (absolute number of referrals, average number of referrals and as percentage of total applications), to generate most workable variable scores.

Second, I included one measurement of organizational status, which might underexpose the importance of status as social judgement. But as Jensen and Roy (2008) argued that status judgements are mostly a selection mechanism and not decisive in decision-making, I assume its impact to be covered sufficiently with one indicator. Lastly, I have not incorporated the effects of mergers, acquisitions and external collaborations on the judgement of the reputation of the applicant. I justify neglecting these effects because of two reasons. Theoretically attention has focused on the impact of reputation on alliance forming (e.g. Stern et al., 2014), but not vice versa, how alliances shape reputation. Besides, it is nearly impossible to track down all acquisitions and collaborations with different institutes and smaller companies. As no well-defined theory is present and data structure would be unreliable, I decided to assume that all previous memory of the CHMP is erased when acquisitions or name chances happened.

As the distribution of my independent variable implicated that it was not suited for the conventional statistical methods for continuous and count data, I performed several analyses using different methods. Linear models, count models and hazard models have been used in similar research (e.g. Carpenter et al., 2010; Kim, 2012). As each of these methods had proven downsides (normal distribution, Poisson distribution, proportional hazards assumption) I decided to incorporate all methods from prior work and focus on differences in outcomes between regression methods. Violating some of the assumptions for these statistical tests implies that coefficients were sometimes harder to interpret (such as for disproportional hazards for some variables). To overcome these problems I performed additional analyses using a cutoff at 200 days of active review time, total review time as dependent variable and introducing clustered standard errors types to provide most complete insights and check for consistency. Since the majority of my regressions provided uniform outcomes regarding which indicators seemed of significant influence, I ensured high inter-method reliability.

The choices made regarding the measurement methods have implications for the outcomes. First of all, I have chosen to include the number of healthy volunteers and patients as measurement of product uncertainty. Despite the theoretical argumentation why the evidence base partly determines product uncertainty, I only found small and insignificant effects. The most likely explanation is that the required clinical data varies a lot per disease area, because drugs to treat diseases with low prevalence cannot obtain large amounts of clinical data. Hence, there is an indication that there should have been included dummy effects per therapeutic class. To correct for this possible effect I performed an additional analysis using dummies per therapeutic class and found no significant differences in outcomes.

The measurements of organizational reputation for being known were replicated from earlier empirical work. However, as the idea of three-dimensional organizational reputation has not been specifically tested in the context of pharmaceuticals, I computed the predictors of being known for something based on the average performances. This indicates that the CHMP remembers an organization's reputation forever. This is contrary to the assumption that I make for mergers and acquisitions, where I assume that reputation disappears after companies merge. Despite this contradiction, I already emphasized on the reason to exclude mergers and acquisitions and persevere that using average scores and proportion to indicate the prior performance of organizations is still more valid than using absolute numbers.

Some important predictors that have neglected in this research are related the impact of the political arena, as Carpenter (2004) has named it. This mostly emphasizes the demand-side of pharmaceuticals, stating that patient advocacy groups and media play a role speeding up drug approvals. Several scholars have given attention to the possible effects of the demand for therapy or media (e.g. Carpenter et al., 2010; Moffit, 2010). I have neglected the demand-side because it has been well-researched and often resulted in finding only marginal or no significant effects on review times. Additionally, I have argued that regulators in pharmaceuticals are assigned to make decisions about

drug quality that patients cannot make by themselves because they behave irrationally regarding the possible risks. This means that regulators are supposed to be indifferent to pressure from the demand-side and therefore justifies neglected the effects of the political arena. Next, another omitted variable from the demand-side was the entry order of the product, which has been proven to affect review times (Stern, 2017). But since products in the EU can be authorized on national level, while not have been approved centrally by the EMA, it is near-impossible to determine when products are the first to treat a particular disease.

The external validity of this research is limited, as it is hard to generalize from this specific context. There has been much research on similar regulatory organizations, mainly for the FDA, however as the FDA is seen as more 'risk-taking' with regard to their choices (Tafuri et al., 2014), it might well be that the effects of organizational reputation are stronger or different and therefore it is hard to compare these outcomes. Moreover, generalizing outcomes of this research towards other industries should be done carefully and primarily can be done for other regulated industries. For example, it has already been proven that organizational reputation plays part in the banking industry (Englert et al., 2018). But the effects of organizational reputation have also been proved for other industries where rating agencies play a part in consumer behavior, such as for the automotive industry (Rao, 1994). Hence, generalizing the results from this research is possible to a certain degree, especially towards context where the social judgement about the organizational reputation is not made by the one who uses the product.

5.5 Future research suggestions

This research can serve as a baseline for several other theoretical and empirical contributions. To begin with, I have tested whether reputation is three-dimensional, as proposed by Lange et al. (2011) and found evidence that it is. More empirical work in varying context focusing on the three-dimensionality of reputation would strengthen the argument by Lange et al. (2011). As a second, this research is based on similar empirical work performed on the FDA. However, since there are still several methodological differences, it does not allow for comparison. Since the FDA is generally said to be more 'risk-taking' (Tafuri et al., 2014), it would be interesting to be able to compare whether the effects of organizational judgements are comparable, or larger than at the CHMP. This would ask for innovative measurements or data gathering of the dimension 'being known for something', as the FDA is less open about for example failed applications.

Thirdly, Carpenter et al. (2012) showed that the administrative deadlines at the FDA are linked to more post marketing safety problems. I found that the CHMP shows similar behavior, approving more than half of all MAAs within the last ten days of product assessment. Following the logic of Carpenter et al. (2012) that deadlines may shape decision quality, it could be expected that the quality of decision at the CHMP is also affected by administrative deadlines. If the CHMP quality of decision is in fact affected by these administrative deadlines, it could mean that products are allowed to the market with higher levels of uncertainty regarding patient safety. Hence, a result of this opportunistic behavior could be that patients use these authorized products assuming they are safe, while the EMA cannot guarantee their safety because there is no complete risk-profile of the product.

A second area of exploration would be to investigate the effects that institutionalization at the EMA has had on active review times. Several institutional scholars such as Meyer and Rowan (1977) and Scott (1995) have laid emphasis how organizations evolve over time and become slower and more bureaucratic as they grow. I found some theoretical indications that safety issues precede new regulations and guidelines, which could lengthen the review process (Carpenter, 2002; Maor, 2011). As I found that active review times have significantly increased since the first safety withdrawal, it could be that institutionalization plays a role in the development of review times. Future research could aim at mapping how institutionalization has affected the EMA since its foundation.

6. Conclusion

This thesis focused on the how organizational reputation affect the social judgement by the CHMP when reviewing MAAs. Using three dimensions of organizational reputation as proposed by Lange et al. (2011) allowed to extend current knowledge on regulatory decision-making under conditions of inherent product uncertainty as is the case for new drug approvals (Carpenter et al., 2010). Previous work focused almost exclusively on two of these dimensions, namely whether applicants are known by the regulator and their generalized favorability. While their findings indicated that organizational reputation does affect the time for decision-making, they have largely overlooked the effect that firm-specific past performances have had. Moreover, they have exclusively focused on the FDA and have not provide insights into decision-making in the European pharmaceutical industry. To provide a full understanding of the effects of organizational reputation in the context of the European pharmaceutical industry, to following research question has been answered: *To what extent do three different types of organizational reputation influence the length of the active review process of marketing authorization applications by the CHMP?*

My results show that social judgements about organizational reputation play a significant role in the time it takes the CHMP to formulate a decision on MAAs when facing product uncertainty. I question the beliefs by Lange et al. (2011) that the familiarity between applicant and regulator yields a non-evaluative judgement, as I found that it significantly extends active review. Considering the effects of organizational reputation, it seems social judgments are mostly derived from the generalized favorability of the applicants, while I also found evident indications that firm-specific past performances are also considered when the CHMP reviews MAAs. Hence, in sum this research has provided empirical evidence that even in industries that greatly rely on scientific information as base of decision-making, the effect of social judgements of the organizational reputation of the applying organization cannot be neglected.

References

Abrams, D. & Sampat, B. (2017). What's the value of patent citations? evidence from pharmaceuticals. Working Paper. Retrieved on 4/4/2019, available at: http://www.law.northwestern.edu/research-faculty/searlecenter/events/innovation/documents/AbramsSampatDrugCites060917.pdf

Aldrich, H. E., & Fiol, C. M. (1994). Fools rush in? The institutional context of industry creation. *Academy of management review*, *19*(4), 645-670.

Alqahtani, S., Seoane-Vazquez, E., Rodriguez-Monguio, R., & Eguale, T. (2015). Priority review drugs approved by the FDA and the EMA: time for international regulatory harmonization of pharmaceuticals?. *Pharmacoepidemiology and drug safety*, *24*(7), 709-715.

Baird, L. G., Banken, R., Eichler, H. G., Kristensen, F. B., Lee, D. K., Lim, J. C. W. & Samaha, D. (2014). Accelerated access to innovative medicines for patients in need. *Clinical Pharmacology & Therapeutics*, *96*(5), 559-571.

Basdeo, D. K., Smith, K. G., Grimm, C. M., Rindova, V. P., & Derfus, P. J. (2006). The impact of market actions on firm reputation. *Strategic Management Journal*, *27*(12), 1205-1219.

Basu, A., Manning, W. G., & Mullahy, J. (2004). Comparing alternative models: log vs Cox proportional hazard?. *Health economics*, *13*(8), 749-765.

Benjamin, B. A., & Podolny, J. M. (1999). Status, quality, and social order in the California wine industry. *Administrative science quarterly*, *44*(3), 563-589.

Beishon, M. A. R. C. (2014). Approval rating: how do the EMA and FDA compare?. *Cancer World*, 12-17.

Bergh, D. D., Connelly, B. L., Ketchen Jr, D. J., & Shannon, L. M. (2014). Signalling theory and equilibrium in strategic management research: An assessment and a research agenda. *Journal of Management Studies*, *51*(8), 1334-1360.

Bitektine, A. (2011). Toward a theory of social judgments of organizations: The case of legitimacy, reputation, and status. *Academy of management review*, *36*(1), 151-179.

Bonardi, J. P., Holburn, G. L., & Vanden Bergh, R. G. (2006). Nonmarket strategy performance: Evidence from US electric utilities. *Academy of Management Journal*, *49*(6), 1209-1228.

Bouvy, J. C., Huinink, L., & De Bruin, M. L. (2016). Benefit-risk reassessment of medicines: a retrospective analysis of all safety-related referral procedures in Europe during 2001–2012. *Pharmacoepidemiology and drug safety*, *25*(9), 1004-1014.

Cameron, A. C., & Windmeijer, F. A. (1996). R-squared measures for count data regression models with applications to health-care utilization. *Journal of Business & Economic Statistics*, *14*(2), 209-220.

Carpenter, D. P. (2002). Groups, the media, agency waiting costs, and FDA drug approval. *American Journal of Political Science*, 490-505.

Carpenter, D. P. (2004). The political economy of FDA drug review: Processing, politics, and lessons for policy. *Health Affairs*, *23*(1), 52-63.

Carpenter, D., Moffitt, S. I., Moore, C. D., Rynbrandt, R. T., Ting, M. M., Yohai, I., & Zucker, E. J. (2010). Early entrant protection in approval regulation: Theory and evidence from FDA drug review. *The Journal of Law, Economics, & Organization*, *26*(3), 515-545.

Cattaneo, M., Meoli, M., & Paleari, S. (2016). Why do universities internationalize? Organizational reputation and legitimacy. In D. Audretsch (eds)., *University Evolution, Entrepreneurial Activity and Regional Competitiveness* (pp. 327-346). Basel: Springer International Publishing.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of management*, *37*(1), 39-67.

Cooper, J. C., & Kovacic, W. E. (2012). Behavioral economics: implications for regulatory behavior. *Journal of Regulatory Economics*, *41*(1), 41-58.

Deephouse, D. L. (1996). Does isomorphism legitimate?. *Academy of management journal*, *39*(4), 1024-1039.

Deephouse, D. L. (2000). Media reputation as a strategic resource: An integration of mass communication and resource-based theories. *Journal of management*, *26*(6), 1091-1112.

Deephouse, D. L., & Carter, S. M. (2005). An examination of differences between organizational legitimacy and organizational reputation. *Journal of management Studies*, *42*(2), 329-360

Deephouse, D. L., & Suchman, M. (2008). Legitimacy in organizational institutionalism. *The Sage handbook of organizational institutionalism*, *49*, 77.

Deephouse, D. L., Bundy, J., Tost, L. P., & Suchman, M. C. (2017). Organizational legitimacy: Six key questions. *The SAGE handbook of organizational institutionalism*, 27-54.

Downing, N. S., Aminawung, J. A., Shah, N. D., Braunstein, J. B., Krumholz, H. M., & Ross, J. S. (2012). Regulatory review of novel therapeutics—comparison of three regulatory agencies. *New England Journal of Medicine*, *366*(24), 2284-2293.

Downing, N. S., Zhang, A. D., & Ross, J. S. (2017). Regulatory Review of New Therapeutic Agents—FDA versus EMA, 2011–2015. *New England Journal of Medicine*, *376*(14), 1386-1387.

Duijnhoven, R. G., Straus, S. M., Raine, J. M., de Boer, A., Hoes, A. W., & De Bruin, M. L. (2013). Number of Patients Studied Prior to Approval of New Medicines: A Database Analysis. *PLoS Medicine*, *10*(3), 1-8.

Eder, J., Sedrani, R., & Wiesmann, C. (2014). The discovery of first-in-class drugs: origins and evolution. *Nature Reviews Drug Discovery*, *13*(8), 577.

Ehmann, F., Amati, M. P., Salmonson, T., Posch, M., Vamvakas, S., Hemmings, R., Eichler, H., G. & Schneider, C. K. (2013). Gatekeepers and enablers: how drug regulators respond to a challenging and changing environment by moving toward a proactive attitude. *Clinical Pharmacology & Therapeutics*, *93*(5), 425-432.

Eichler, H. G., Pignatti, F., Flamion, B., Leufkens, H., & Breckenridge, A. (2008). Balancing early market access to new drugs with the need for benefit/risk data: a mounting dilemma. *Nature Reviews Drug Discovery*, *7*(10), 818-826

Elsäßer, A., Regnstrom, J., Vetter, T., Koenig, F., Hemmings, R. J., Greco, M., Papaluca-Amati, M. & Posch, M. (2014). Adaptive clinical trial designs for European marketing authorization: a survey of scientific advice letters from the European Medicines Agency. *Trials*, *15*(1), 383.

EMA. (2019). European Medicines Agency pre-authorisation procedural advice for users of the centralised procedure. Retrieved on 7/4/2019, available at: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/european-medicines-agency-pre-authorisation-procedural-advice-users-centralised-procedure_en-0.pdf

Englert, M. R., Koch, C., & Wüstemann, J. (2018). The effects of financial crisis on the organizational reputation of banks: An empirical analysis of newspaper articles. *Business & Society*, 1-35.

European Commission. (2001). 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. Retrieved on 7/4/2019, available at: https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/dir_2001_83_consol_2012/dir_2001_83_cons_2012_en.pdf

Grabowski, H. (2002). Patents, innovation and access to new pharmaceuticals. *Journal of International Economic Law*, *5*(4), 849-860.

Harford, J., Klasa, S., & Walcott, N. (2009). Do firms have leverage targets? Evidence from acquisitions. *Journal of Financial Economics*, *93*(1), 1-14.

Hauray, B. (2017). From regulatory knowledge to regulatory decisions: the European evaluation of medicines. *Minerva*, *55*(2), 187-208.

Hoekman, J., & Boon, W. (2019). Changing standards for drug approval: A longitudinal analysis of conditional marketing authorisation in the European Union. *Social Science & Medicine*, *222*, 76-83.

Hofer, M. P., Hedman, H., Mavris, M., Koenig, F., Vetter, T., Posch, M., Vamvakas, F., Regnstrom, J. & Aarum, S. (2018). Marketing authorisation of orphan medicines in Europe from 2000 to 2013. *Drug discovery today*, *23*(2), 424-433.

Hu, Y., McNamara, P., & McLoughlin, D. (2015). Outbound open innovation in bio-pharmaceutical out-licensing. *Technovation*, *35*, 46-58.

Jayasundara, K., Hollis, A., Krahn, M., Mamdani, M., Hoch, J. S., & Grootendorst, P. (2019). Estimating the clinical cost of drug development for orphan versus non-orphan drugs. *Orphanet journal of rare diseases*, *14*(1), 1-10

Jensen, M., & Roy, A. (2008). Staging exchange partner choices: When do status and reputation matter?. *Academy of Management Journal*, *51*(3), 495-516.

Jensen, M., Kim, B. K., & Kim, H. (2011). The importance of status in markets: A market identity perspective. In J. L. Pearce (Ed.)*, Status in management and organizations* (pp. 87-117). Cambridge: Cambridge University Press.

Kashoki, M., Hanaizi, Z., Yordanova, S., Veselý, R., Bouygues, C., Llinares, J., & Kweder, S. L. (2019). A Comparison of EMA and FDA Decisions for New Drug Marketing Applications 2014–2016: Concordance, Discordance, and Why. *Clinical Pharmacology & Therapeutics.* 1-8

Kaitin, K. I. (2010). Deconstructing the drug development process: the new face of innovation. *Clinical Pharmacology & Therapeutics*, *87*(3), 356-361.

Kim, J. W. (2012). Arbiter of science: Institutionalization and status effects in FDA drug review 1990–2004. *Strategic Organization*, *10*(2), 128-157

King, B. G., & Whetten, D. A. (2008). Rethinking the relationship between reputation and legitimacy: A social actor conceptualization. *Corporate Reputation Review*, *11*(3), 192-207.

Kyle, M. K. (2006). The role of firm characteristics in pharmaceutical product launches. *The RAND Journal of Economics*, *37*(3), 602-618.

Lange, D., Lee, P. M., & Dai, Y. (2011). Organizational reputation: A review. *Journal of management*, *37*(1), 153-184.

Liberti, L., Breckenridge, A., Hoekman, J., McAuslane, N., Stolk, P., & Leufkens, H. (2017). Factors related to drug approvals: predictors of outcome?. *Drug discovery today*, *22*(6), 937-946.

MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of econometrics, 29(3),* 305-325.

Makuch, R. W., & Shi, R. (2014). Comparison of drug approvals in europe versus the United States: An analysis of discrepancies between drug products reviewed by EMA and FDA. *Therapeutic innovation & regulatory science*, *48*(3), 362-366.

Maor, M. (2011). Organizational reputations and the observability of public warnings in 10 pharmaceutical markets. *Governance*, *24*(3), 557-582.

McNaughton, R., Huet, G., & Shakir, S. (2014). An investigation into drug products withdrawn from the EU market between 2002 and 2011 for safety reasons and the evidence used to support the decision-making. *BMJ open*, *4*(1), 1-6

Moffitt, S. L. (2010). Promoting agency reputation through public advice: Advisory committee use in the FDA. *The Journal of Politics*, *72*(3), 880-893.

Morgan, S., Grootendorst, P., Lexchin, J., Cunningham, C., & Greyson, D. (2011). The cost of drug development: a systematic review. *Health policy*, *100*(1), 4-17.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. American journal of sociology, 83(2), 340-363.

Motola, D., De Ponti, F., Poluzzi, E., Martini, N., Rossi, P., Silvani, M. C. & Montanaro, N. (2006). An update on the first decade of the European centralized procedure: how many innovative drugs?. *British journal of clinical pharmacology*, *62*(5), 610-616.

Mullins, C. D., Montgomery, R., & Tunis, S. (2010). Uncertainty in assessing value of oncology treatments. *The oncologist*, *15*(S1), 58-64.

Noll, R., G. (1976). *Government regulation and behavior: A multidisciplinary survey and synthesis.* California: California Institute for Technology.

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & quantity*, *41*(5), 673-690.

Olson, M. K. (1997). Firm characteristics and the speed of FDA approval. *Journal of economics & management strategy*, *6*(1), 377-401.

Olson, M. K. (2004). Managing delegation in the FDA: Reducing delay in new-drug review. *Journal of Health Politics, Policy and Law*, *29*(3), 397-430.

Patterson, K., Cavazos, D., & Washington, M. (2014). It does matter how you get to the top: Differentiating status from reputation. *Administrative Sciences*, *4*(2), 73-86.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, *22*(1), 435-480.

Petkova, A. P., Wadhwa, A., Yao, X., & Jain, S. (2014). Reputation and decision making under ambiguity: a study of US venture capital firms' investments in the emerging clean energy sector. *Academy of Management Journal*, *57*(2), 422-448.

Phillips, D. J., & Zuckerman, E. W. (2001). Middle-status conformity: Theoretical restatement and empirical demonstration in two markets. *American Journal of Sociology*, *107*(2), 379-429.

Piazza, A., & Castellucci, F. (2014). Status in organization and management theory. *Journal of Management*, *40*(1), 287-315.

Podolny, J. M. (1993). A status-based model of market competition. *American journal of sociology*, *98*(4), 829-872.

Polidoro Jr, F., & Theeke, M. (2012). Getting competition down to a science: The effects of technological competition on firms' scientific publications. *Organization Science*, *23*(4), 1135-1153.

Putzeist, M. (2013). *Marketing authorisation of new medicines in the EU: Towards evidence-based improvement*. Amsterdam: Off Page

Rao, H. (1994). The social construction of reputation: Certification contests, legitimation, and the survival of organizations in the American automobile industry: 1895–1912. *Strategic management journal*, *15*(S1), 29-44.

Regnström, J., Koenig, F., Aronsson, B., Reimer, T., Svendsen, K., Tsigkos, S., Flamion, B., Eichler, H.G.. & Vamvakas, S. (2010). Factors associated with success of market authorisation applications for pharmaceutical drugs submitted to the European Medicines Agency. *European journal of clinical pharmacology*, *66*(1), 39.

Rhee, M., & Haunschild, P. R. (2006). The liability of good reputation: A study of product recalls in the US automobile industry. *Organization Science*, *17*(1), 101-117.

Rindova, V. P., Williamson, I. O., Petkova, A. P., & Sever, J. M. (2005). Being good or being known: An empirical examination of the dimensions, antecedents, and consequences of organizational reputation. *Academy of management journal*, *48*(6), 1033-1049.

Rindova, V. P., Pollock, T. G., & Hayward, M. L. (2006). Celebrity firms: The social construction of market popularity. *Academy of management review*, *31*(1), 50-71.

Roberts, S. A., Allen, J. D., & Sigal, E. V. (2011). Despite criticism of the FDA review process, new cancer drugs reach patients sooner in the United States than in Europe. *Health affairs*, *30*(7), 1375-1381.

Ruef, M., & Scott, W. R. (1998). A multidimensional model of organizational legitimacy: Hospital survival in changing institutional environments. *Administrative science quarterly*, 877-904.

Scott, W. R. (1995). *Institutions and organizations.* Thousand Oaks, California: SAGE Publications

Schober, P., & Vetter, T. R. (2018). Survival analysis and interpretation of time-to-event data: The tortoise and the hare. *Anesthesia and analgesia*, *127*(3), 792.

Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika*, *69*(1), 239-241.

Shah, R. R., Roberts, S. A., & Shah, D. R. (2013). A fresh perspective on comparing the FDA and the CHMP/EMA: approval of antineoplastic tyrosine kinase inhibitors. *British journal of clinical pharmacology*, *76*(3), 396-411.

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics, 87*(3), 355-374.

Stern, A. D. (2017). Innovation under regulatory uncertainty: evidence from medical technology. *Journal of public economics*, *145*, 181-200.

Stern, I., Dukerich, J. M., & Zajac, E. (2014). Unmixed signals: How reputation and status affect alliance formation. *Strategic Management Journal*, *35*(4), 512-531.

Sorenson, O. (2014). Status and reputation: Synonyms or separate concepts?. *Strategic Organization*, *12*(1), 62-69.

Stuart, T. E., Hoang, H., & Hybels, R. C. (1999). Interorganizational endorsements and the performance of entrepreneurial ventures. *Administrative science quarterly*, *44*(2), 315-349.

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of management review*, *20*(3), 571-610.

Tafuri, G., Stolk, P., Trotta, F., Putzeist, M., Leufkens, H. G., Laing, R. O., & De Allegri, M. (2014). How do the EMA and FDA decide which anticancer drugs make it to the market? A comparative qualitative study on decision makers' views. *Annals of oncology*, *25*(1), 265-269.

Taylor, R. (1990). Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography*, *6*(1), 35-39.

Therneau, T., Crowson, C., & Atkinson, E. (2019) Using Time Dependent Covariates and Time Dependent Coefficients in the Cox Model. Retrieved on 19/9/2019, available at: https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf

Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of financial Economics*, *99*(1), 1-10.

Trusheim, M. R., Aitken, M. L., & Berndt, E. R. (2010). Characterizing Markets for Biopharmaceutical Innovations: Do Biologics Differ from Small Molecules?. *Frontiers in Health Policy Research*, *13*(1), 1-43.

Ver Hoef, J. M., & Boveng, P. L. (2007). Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data?. *Ecology*, *88*(11), 2766-2772.

Washington, M., & Zajac, E. J. (2005). Status evolution and competition: Theory and evidence. *Academy of Management Journal*, *48*(2), 282-296.

White, K. J. (1992). The Durbin-Watson test for autocorrelation in nonlinear models. *The Review of Economics and Statistics,* 370-373.

Wood, G. R. (2002). Assessing goodness of fit for Poisson and negative binomial models with low mean. *Communications in Statistics - Theory and Methods*, *31(11),* 1977-2001

Wooldridge, J. (2007). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

Green, M., & Moore Jr, B. (1972). Winter's Discontent: Market Failure and Consumer Welfare. *Yale LJ*, 82, 903.

Appendix A – Search terms in Web of Science

Table 15 provides the search terms used to find the organization in Web of Science and obtain the average number of citations on scientific publications published by that organization. Companies of which the name is bold were found via the "organization enhanced" option, also including publications by foreign subsidiaries and research facilities.

Table 15: Search terms to find average number of citations in Web of Science

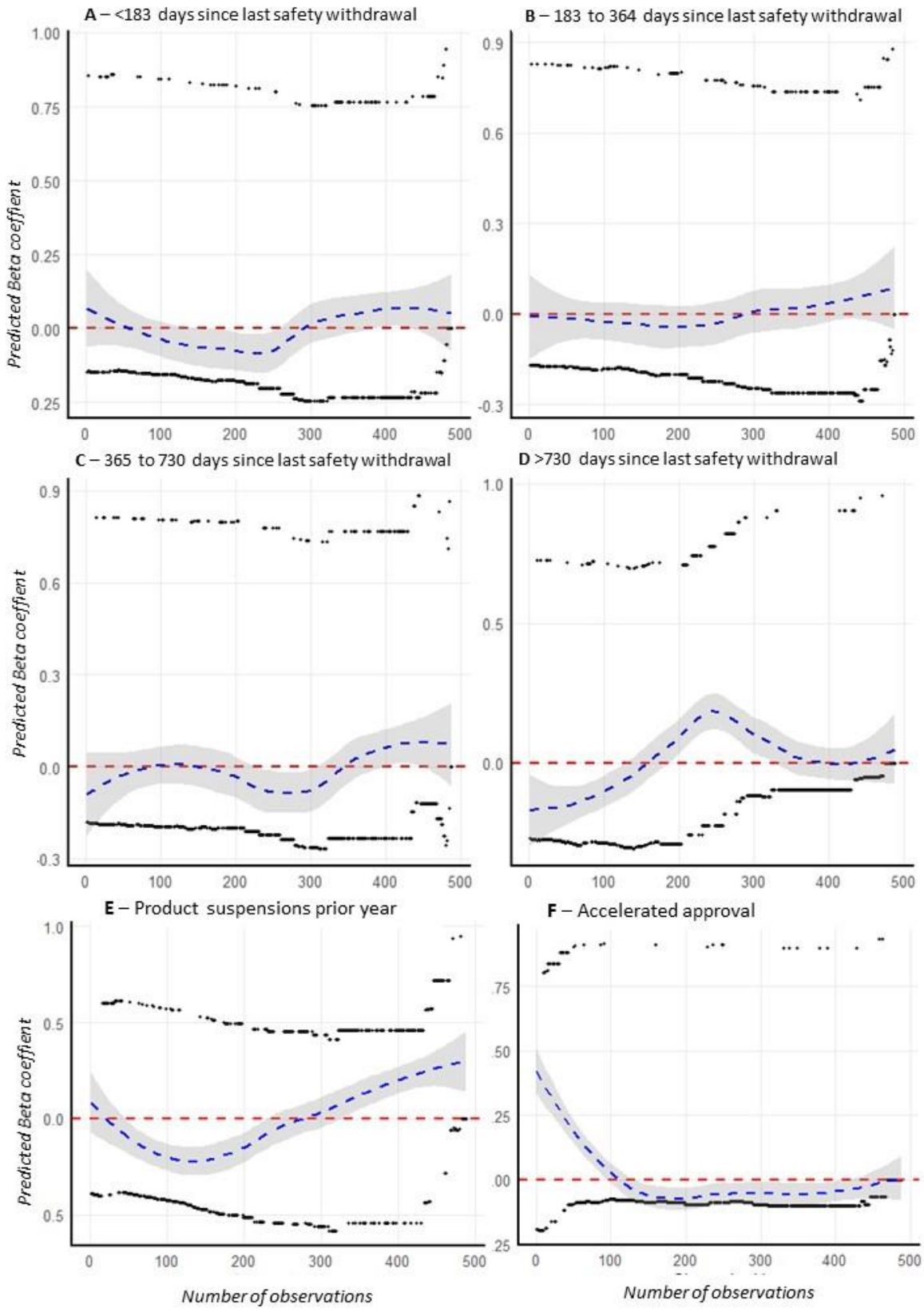
	Applicant name	Search term		Applicant name	Search term
1	**3m**	3M	83	nycomed	Nycomed
2	**abbott laboratories**	ABBOTT GMBH & CO. KG OR Abbott Laboratories OR Abbott Turkiye OR Abbott Vascular OR AbbVie	84	octapharma	Octapharma
3	abraxis biosciences	Abraxis	85	omrix biopharmaceuticals	Omrix
4	actavis	Actavis	86	orion	Orion
5	**actelion**	Actelion	87	orphan europe	Orphan
6	aegerion	Aegerion	88	**otsuka**	Otsuka
7	agouron	Agouron	89	pharmion	Pharmion
8	alexion	Alexion	90	**pfizer**	Pfizer
9	alexza	Alexza	91	**pharmacia**	Pharmacia
10	**allergan**	Allergan	92	pharming	Pharming
11	**almirall**	Almirall	93	pierre fabre	(Pierre AND Fabre)
12	amag pharmaceuticals	Amag	94	gedeon richter	(Gedeon AND Richter)
13	**amgen**	Amgen	95	**procter and gamble**	(Procter & Gamble)
14	**apotex**	Apotex	96	prostrakan	Prostrakan
15	ares serono	Ares AND Serono	97	ratiopharm	Ratiopha*
16	ariad	Ariad	98	recordati	Recordati
17	asta medica	Asta Medica	99	regeneron	Regeneron
18	**astellas**	Astellas	100	**roche**	Roche
19	**astrazeneca**	Astrazeneca	101	**novartis**	Novartis
20	aventis	Aventis	102	**sanofi aventis**	(Sanofi AND Aventis)
21	**sanofi**	Sanofi	103	savient	Savient
22	axcan	Axcan	104	schering ag	Schering AG
23	**baxter**	Baxter	105	**schering plough**	(Schering AND Plough)
24	**bayer**	Bayer AG OR Bayer Healthcare Pharmaceuticals OR Bayer Turkey	106	schwarz	Schwarz
25	behringwerke	Behringwerke	107	**shire**	Shire
26	**bial**	Bial	108	sigma tau	(Sigma AND Tau)
27	biocodex	Biocodex	109	sonus pharma	Sonus
28	bioenvision	Bioenvis	110	sparkle	Sparkle pharm*
29	**biogen idec**	Biogen	111	**takeda**	Takeda
30	biomarin	Biomarin	112	tercica	Tercica
31	biotest	Biotest	113	**teva pharmaceuticals**	Teva pharm*

	Applicant name	Search term		Applicant name	Search term
32	**boehringer ingelheim**	Boehringer	114	the medicines company	(Medicines AND Co)
33	**bracco**	Bracco	115	thrombogenics	Thrombogenics
34	**bristol myers squibb**	Bristol-Myers Squibb OR Bristol Myers Squibb - Turkey	116	tigenix	Tigenix
35	cangene	Cangene	117	transkaryotic therapies	Transkaryo*
36	**celgene**	Celgene	118	tmc pharma	tmc
37	**johnson and johnson**	Johnson & Johnson OR Johnson & Johnson USA	119	topotarget	Topotarget
38	cephalon	Cephalon	120	triangle pharmaceuticals	(Triangle AND Pharm*)
39	**chiesi**	Chiesi	121	tyco	Tyco
40	chiron	Chrion	122	**ucb**	UCB
41	clinuvel	Clinuvel	123	united therapeutics	(United AND therapeutics)
42	croma	Croma	124	vanda pharmaceuticals	(Vanda AND pharma*)
43	**csl**	CSL	125	vertex pharmaceuticals	(Vertex AND pharma*)
44	cti biopharma	Cti | cti AND bio*	126	vifor pharma	Vifor
45	cv therapeutics	cv therapeutic	127	**yamanouchi**	Yamanouchi
46	dendreon	Dendreon	128	viiv healthcare	Viiv
47	eisai	Eisai	129	viropharma	Viroph*
48	elan pharmaceuticals	(Elan AND Pharm*)	130	vivus	Vivus
49	**eli lilly**	Lilly	131	glaxo wellcome	(Glaxo AND Wellcome)
50	encysive	Encysive	132	rhone poulenc	(Rhone AND poulenc)
51	eusa pharma	Eusa	133	immunomedics	Immunomed*
52	ferring	Ferring	134	ciba geigy	Ciba
53	fgk	Fgk	135	watson pharmaceuticals	(Watson AND pharm*)
54	forest laboratories	Forest AND lab	136	**daiichi sankyo**	Daiichi
55	fresenius	Fresenius	137	**alcon laboratories**	Alcon
56	fujisawa	Fujisawa	138	sanofi synthelabo	Synthelabo
57	galderma	Galderna	139	**merck kgaa**	Merck Kgaa
58	wyeth	Wyeth	140	aptalis	Aptalis
59	**genzyme**	Genzyme	141	basilea pharmaceutica	Basilea
60	**gilead**	Gilead Sciences	142	genetics institute	(Genetic* AND Inst*)
61	**glaxosmithkline**	Glaxosmithkline	143	birken	Birken
62	**lundbeck**	Lundbeck	144	cubist pharmaceuticals	Cubist
63	helsinn	Helsinn	145	durata therapeutics	Durata
64	idm	Idm	146	horizon pharma	(Horizon AND pharm*)
65	grifols	Grifols	147	impax laboratories	Impax
66	intermune	Intermune	148	keryx biopharmaceuticals	Keryx
67	ipsen	Ipsen	149	omeros	Omeros
68	jerini	Jerini	150	santen	Saneten
69	hra	Hra	151	shionogi	Shionogi
70	leo pharma	(Leo AND pharm*)	152	synageva biopharma	Synageva
71	**servier**	Servier	153	zambon pharma	Zambon AND pharma

	Applicant name	Search term		Applicant name	Search term
72	**ligand**	Ligan	154	bioton	Bioton
73	mallinckrodt	Mallinckrodt	155	**sanofi bristol myers squibb**	Average of sanofi and BMS
74	medac	Medac	156	**astrazeneca bristol myers squibb**	Average of astrazeneca and BMS
75	**merck and co**	Merck & Company	157	**pfizer bristol myers squibb**	Average of pfizer and BMS
76	merz	Merz	158	amicus	Amicus
77	mitsubishi	Mitsubishi pharma	159	bio products laboratory	Bio products laboratory
78	movetis	Movetis	160	bioprojet	Bioprojet
79	organon	Organon	161	shield therapeutics	Shield
80	neurim pharmaceuticals	Neurim	162	molmed	Molmed
81	neurogesx	Neuroges*	163	intercept pharmaceuticals	Intercept
82	**novo nordisk**	Nordisk	164	exelixis	Exelexis

Appendix B – Schoenfeld's residuals

Figure 8A-F: Schoenfeld residuals plots of indicators with significantly changing beta-coefficients



Note: Observation 100 = 181 days of review; observation 200 = 201 days; observation 300 = 208 days; observation 400 = 210 days

Appendix C – Variance Inflation Factor scores

Table 16 provides the variance inflation factors of all predictor variables, computed to detect multicollinearity when. The variance inflation factors were calculated after simultaneous introduction to a Poisson regression.

Table 16: Predictors and variance inflation factors

Predictor	VIF
(U) Complexity	1.990
(U) Evidence base	1.256
(U) Approval history	1.238
(BK) Previous approvals	2.557
(BK) Current applications	1.366
(BKS) Average number of product referrals	1.701
(BKS) Firm related safety withdrawals	1.340
(BKS) Voluntary withdrawals	1.508
(BKS) Failed applications	1.242
(GF) Number of product suspensions prior year	1.500
(GF) Time since last safety withdrawal	1.951
(GF) Specialization	1.265
Status	1.306
Age	1.939
Sales	2.686
Applied for accelerated assessment	1.206
Origin	1.118

(U) = variables measuring product uncertainty, (BK) = variables measuring being known, (BKS) = variables measuring being known for something; (GF) = variables measuring generalized favorability

Appendix D – Outcomes of quasi-Poisson regressions

Table 17 gives the regression outcomes of the quasi-Poisson regressions when using active review time as the dependent variable.

Table 17: Coefficients, standard errors and model fit of quasi-Poisson regressions

	Dependent variable: Active review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(U) Product complexity (ln)	**0.040*** **	**0.032** **	**0.037*** **	**0.031** **	**0.039*** **	**0.038*** **	**0.038*** **	**0.037*** **	**0.035*** **	0.014	0.004	**0.036*** **	-0.010	-0.012
	(0.010)	**(0.011)**	**(0.010)**	**(0.011)**	**(0.010)**	**(0.010)**	**(0.010)**	**(0.010)**	**(0.010)**	(0.011)	(0.011)	**(0.010)**	(0.012)	(0.012)
(U) Evidence base (ln)	-0.002	-0.0001	-0.001	0.001	-0.002	-0.002	-0.002	-0.002	-0.002	0.001	0.002	-0.001	0.003	0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
(U) <183 days since FDA approval	**-0.033** **	**-0.035** **	**-0.035** **	**-0.036** **	**-0.034** **	**-0.034** **	**-0.034** **	**-0.034** **	**-0.034** **	**-0.027***	**-0.029***	**-0.031***	**-0.024***	**-0.025***
	(0.012)	**(0.012)**	**(0.012)**	**(0.012)**	**(0.012)**	**(0.012)**	**(0.012)**	**(0.012)**	**(0.012)**	**(0.012)**	**(0.012)**	**(0.012)**	**(0.012)**	**(0.012)**
(U) 183 to 364 days since FDA approval	0.019	0.018	0.017	0.016	0.018	0.016	0.018	0.017	0.014	0.022	0.024	0.022	0.026+	0.023
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)
(U) >365 days since FDA approval	0.0002	0.0003	-0.0003	-0.0002	-0.0002	-0.0001	-0.001	-0.001	-0.001	0.008	0.002	0.0002	0.007	0.006
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.014)	(0.015)	(0.014)	(0.014)
(BK) Past approvals (ln)		**0.013***		0.010										-0.004
		(0.006)		(0.006)										(0.007)
(BK) Current applications (ln)			**0.022***	0.018+										**0.022***
			(0.010)	(0.010)										**(0.010)**
(BKS) Product referrals					0.021				-0.060					-0.056
					(0.047)				(0.058)					(0.056)
(BKS) Involved in safety withdrawal						0.037+			0.039+					0.030
						(0.021)			(0.022)					(0.022)
(BKS) Voluntary withdrawals							0.055		0.070					0.057
							(0.047)		(0.055)					(0.054)
(BKS) Failed applications								0.079+	0.084+					0.018
								(0.047)	(0.049)					(0.048)
(GF) Product suspension prior year										**0.079*** **			**0.057*** **	**0.056*** **
										(0.016)			**(0.016)**	**(0.017)**
(GF) <183 days since last safety withdrawal											**0.094*** **		**0.073*** **	**0.076*** **
											(0.017)		**(0.018)**	**(0.018)**

Table 17: Coefficients, standard errors and model fit of quasi-Poisson regressions

	Dependent variable: Active review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(GF) 183 to 364 days since last safety withdrawal											**0.085*** **		**0.070*** **	**0.069*** **
											(0.016)		**(0.017)**	**(0.017)**
(GF) 365 to 730 days since last safety withdrawal											**0.091*** **		**0.080*** **	**0.080*** **
											(0.017)		**(0.017)**	**(0.017)**
(GF) >730 days since last safety withdrawal											**0.080*** **		**0.064*** **	**0.065*** **
											(0.016)		**(0.017)**	**(0.017)**
(GF) Specialization												**0.051***	0.020	0.018
												(0.022)	(0.022)	(0.022)
Status (ln)	-0.004	-0.005	-0.005	-0.005	-0.004	-0.004	-0.003	-0.003	-0.002	-0.004	-0.002	-0.004	-0.002	-0.002
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Age (ln)	-0.002	-0.004	-0.003	-0.005	-0.002	-0.003	-0.001	-0.001	0.0004	-0.001	-0.00004	0.002	0.002	0.002
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Sales (ln)	-0.0004	-0.002	-0.002	-0.003	-0.0005	-0.001	-0.001	-0.001	-0.002	-0.0002	-0.0003	-0.0003	-0.00005	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Applied for AA	**-0.169*** **	**-0.170*** **	**-0.164*** **	**-0.165*** **	**-0.169*** **	**-0.169*** **	**-0.168*** **	**-0.169*** **	**-0.169*** **	**-0.175*** **	**-0.168*** **	**-0.167*** **	**-0.172*** **	**-0.166*** **
	(0.019)	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.018)**	**(0.018)**	**(0.019)**	**(0.018)**	**(0.018)**
Origin	-0.004	-0.006	-0.007	-0.008	-0.004	-0.004	-0.003	-0.005	-0.004	-0.005	-0.002	-0.003	-0.003	-0.004
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Constant	**5.045*** **	**5.105*** **	**5.066*** **	**5.111*** **	**5.049*** **	**5.063*** **	**5.048*** **	**5.063*** **	**5.074*** **	**5.186*** **	**5.195*** **	**5.008*** **	**5.255*** **	**5.268*** **
	(0.076)	**(0.081)**	**(0.076)**	**(0.081)**	**(0.077)**	**(0.076)**	**(0.076)**	**(0.076)**	**(0.077)**	**(0.079)**	**(0.078)**	**(0.077)**	**(0.083)**	**(0.086)**
Observations	488	488	488	488	488	488	488	488	488	488	488	488	488	488
Dispersion	**2.126*** **	**2.116*** **	**2.109*** **	**2.103*** **	**2.125*** **	**2.114*** **	**2.122*** **	**2.114*** **	**2.098*** **	**2.040*** **	**1.994*** **	**2.109*** **	**1.949*** **	**1.920*** **
R^2	0.189	0.196	0.197	0.201	0.189	0.193	0.191	0.193	0.200	0.230	0.252	0.197	0.274	0.286

*Note: +p<0.1; *p<0.05; **p<0.01; ***p<0.001; (U) = variables measuring product uncertainty, (BK) = variables measuring being known, (BKS) = variables measuring being known for something; (GF) = variables measuring generalized favorability*

Appendix E – Proportional hazards regressions with total review time as dependent variable

Table 18 provides the regression outcomes of the proportional hazards regressions, using the total review time as the dependent variable.

Table 18: Coefficients, standard errors and model fit of proportional hazards regressions

	Dependent variable: Total review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(U) Product complexity (ln)	-0.141	-0.143	-0.107	-0.139	-0.156	-0.123	-0.140	-0.121	-0.125	-0.051	0.074	-0.142	0.070	0.081
	(0.102)	(0.107)	(0.103)	(0.107)	(0.104)	(0.104)	(0.102)	(0.103)	(0.106)	(0.112)	(0.116)	(0.104)	(0.121)	(0.125)
(U) Evidence base (ln)	0.038	0.039	0.011	0.016	0.041	0.039	0.039	0.034	0.040	0.024	0.017	0.039	0.019	0.003
	(0.039)	(0.040)	(0.040)	(0.040)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)	(0.040)	(0.039)	(0.041)	(0.042)
(U) <183 days since FDA approval	**0.384****	**0.384****	**0.395****	**0.387****	**0.374****	**0.391****	**0.388****	**0.399****	**0.400****	**0.379****	**0.373****	**0.384****	**0.377****	**0.381****
	(0.122)	**(0.122)**	**(0.122)**	**(0.122)**	**(0.123)**	**(0.122)**	**(0.123)**	**(0.122)**	**(0.123)**	**(0.122)**	**(0.123)**	**(0.122)**	**(0.123)**	**(0.125)**
(U) 183 to 364 days since FDA approval	-0.205	-0.206	-0.191	-0.199	-0.217	-0.190	-0.202	-0.194	-0.190	-0.214	-0.255	-0.204	-0.242	-0.199
	(0.157)	(0.157)	(0.156)	(0.157)	(0.157)	(0.157)	(0.157)	(0.157)	(0.158)	(0.156)	(0.159)	(0.158)	(0.160)	(0.161)
(U) >365 days since FDA approval	**-0.305***	**-0.305***	**-0.324***	**-0.324***	**-0.313***	**-0.305***	**-0.302***	**-0.306***	**-0.315***	**-0.341***	**-0.327***	**-0.305***	**-0.330***	**-0.364***
	(0.145)	**(0.145)**	**(0.146)**	**(0.146)**	**(0.146)**	**(0.145)**	**(0.146)**	**(0.145)**	**(0.145)**	**(0.146)**	**(0.147)**	**(0.145)**	**(0.149)**	**(0.149)**
(BK) Past approvals		0.0004		0.008										**0.018***
		(0.007)		(0.007)										**(0.008)**
(BK) Current applications			**-0.096****	**-0.109****										**-0.130*****
			(0.036)	**(0.038)**										**(0.039)**
(BKS) Product referrals					0.393				1.002+					1.058+
					(0.450)				(0.562)					(0.554)
(BKS) Involved in safety withdrawal						-0.182			-0.247					**-0.545***
						(0.210)			(0.217)					**(0.258)**
(BKS) Voluntary withdrawals							-0.146		-0.523					-0.428
							(0.527)		(0.620)					(0.630)
(BKS) Failed applications								-0.607	-0.792					-0.196
								(0.476)	(0.494)					(0.503)
(GF) Product suspension prior year										**-0.204***			-0.021	-0.094
										(0.103)			(0.114)	(0.118)
(GF) <183 days since last safety withdrawal										**-0.051**	**-0.742*****		**-0.754*****	**-0.759*****
											(0.162)		**(0.174)**	**(0.176)**
(GF) 183 to 364 days since last safety withdrawal											**-0.761*****		**-0.776*****	**-0.744*****
											(0.159)		**(0.169)**	**(0.169)**
(GF) 365 to 730 days since last safety withdrawal											**-0.488****		**-0.504****	**-0.461****
											(0.167)		**(0.170)**	**(0.174)**

Table 18 (continued): Coefficients, standard errors and model fit of proportional hazards regressions

	Dependent variable: Total review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(GF) >730 days since last safety withdrawal											**-0.415***		**-0.419***	**-0.370***
											(0.161)		**(0.164)**	**(0.168)**
(GF) Specialization												0.009	0.210	0.259
												(0.206)	(0.216)	(0.219)
Status (ln)	0.049	0.049	0.052	0.051	0.057	0.049	0.047	0.041	0.053	0.048	0.054	0.049	0.055	0.071
	(0.063)	(0.063)	(0.064)	(0.063)	(0.064)	(0.063)	(0.064)	(0.063)	(0.064)	(0.063)	(0.064)	(0.063)	(0.064)	(0.064)
Age (ln)	-0.003	-0.004	0.009	-0.002	-0.006	0.002	-0.008	-0.016	-0.035	-0.007	0.013	-0.003	0.026	0.015
	(0.051)	(0.053)	(0.051)	(0.052)	(0.051)	(0.052)	(0.054)	(0.052)	(0.057)	(0.051)	(0.051)	(0.053)	(0.053)	(0.060)
Sales (ln)	**0.072*** **	**0.072*** **	**0.090*** **	**0.087*** **	**0.072*** **	**0.073*** **	**0.074*** **	**0.079*** **	**0.086*** **	**0.072*** **	**0.071*** **	**0.072*** **	**0.071*** **	**0.090*** **
	(0.021)	**(0.022)**	**(0.023)**	**(0.023)**	**(0.021)**	**(0.021)**	**(0.022)**	**(0.022)**	**(0.023)**	**(0.021)**	**(0.021)**	**(0.021)**	**(0.021)**	**(0.024)**
Applied for AA	**1.815*** **	**1.814*** **	**1.767*** **	**1.744*** **	**1.819*** **	**1.840*** **	**1.812*** **	**1.814*** **	**1.848*** **	**1.848*** **	**1.873*** **	**1.816*** **	**1.904*** **	**1.909*** **
	(0.177)	**(0.177)**	**(0.177)**	**(0.179)**	**(0.177)**	**(0.179)**	**(0.177)**	**(0.177)**	**(0.180)**	**(0.178)**	**(0.179)**	**(0.179)**	**(0.182)**	**(0.189)**
Origin	0.025	0.024	0.049	0.031	0.023	0.032	0.021	0.039	0.031	0.033	0.013	0.025	0.014	0.003
	(0.094)	(0.096)	(0.095)	(0.096)	(0.094)	(0.094)	(0.095)	(0.094)	(0.096)	(0.094)	(0.094)	(0.094)	(0.094)	(0.098)
Observations	488	488	488	488	488	488	488	488	488	488	488	488	488	488
R^2	0.247	0.247	0.259	0.261	0.248	0.248	0.247	0.250	0.255	0.253	0.289	0.247	0.291	0.315
Log Likelihood	-2467.702	-2467.700	-2463.837	-2463.227	-2467.351	-2467.313	-2467.663	-2466.850	-2465.038	-2465.753	-2453.525	-2467.701	-2453.037	-2444.435

*Note: +p<0.1; *p<0.05; **p<0.01; ***p<0.001; (U) = variables measuring product uncertainty, (BK) = variables measuring being known, (BKS) = variables measuring being known for something; (GF) = variables measuring generalized favorability*

Appendix F – Negative binomial regressions with clustered standard errors

Table 19 gives the regression outcomes of the negative binomial regression after using standard errors that were clustered by applicant.

Table 19: Coefficients and standard errors of negative binomial regression using standard errors clustered by applicant

	Independent variable: Active review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(U) Product complexity (ln)	**0.040*** **	**0.031** **	**0.037*** **	**0.031** **	**0.039*** **	**0.037*** **	**0.038*** **	**0.036*** **	**0.034*** **	0.014	0.003	**0.036*** **	-0.011	-0.013
	(0.011)	**(0.011)**	**(0.011)**	**(0.011)**	**(0.011)**	**(0.010)**	**(0.011)**	**(0.010)**	**(0.010)**	(0.013)	(0.011)	**(0.011)**	(0.013)	(0.013)
(U) Evidence base (ln)	-0.002	0.0001	-0.0003	0.001	-0.002	-0.001	-0.002	-0.002	-0.002	0.001	0.002	-0.001	0.004	0.005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
(U) <183 days since FDA approval	**-0.033***	**-0.035***	**-0.035***	**-0.036***	**-0.034***	**-0.034***	**-0.034***	**-0.034***	**-0.034***	**-0.027***	**-0.029***	**-0.031***	-0.024+	-0.026+
	(0.014)	**(0.014)**	**(0.014)**	**(0.014)**	**(0.014)**	**(0.014)**	**(0.014)**	**(0.014)**	**(0.014)**	**(0.014)**	**(0.013)**	**(0.014)**	(0.013)	(0.014)
(U) 183 to 364 days since FDA approval	0.020+	0.019	0.018	0.017	0.019	0.017	0.019	0.018	0.015	0.023+	**0.025***	0.024+	**0.027***	0.023+
	(0.012)	(0.012)	(0.012)	(0.011)	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	**(0.012)**	(0.012)	**(0.013)**	(0.012)
(U) >365 days since FDA approval	0.0003	0.0004	-0.0001	0.00001	-0.0002	-0.00005	-0.001	-0.001	-0.001	0.008	0.002	0.0003	0.007	0.006
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)
(BK) Past approvals (ln)		0.013+		0.010										-0.004
		(0.007)		(0.007)										(0.005)
(BK) Current applications (ln)			**0.022***	0.019+										**0.022***
			(0.009)	(0.010)										**(0.009)**
(BKS) Product referrals					0.021				-0.063					-0.059
					(0.045)				(0.065)					(0.061)
(BKS) Involved in safety withdrawal						0.039+			**0.041***					0.032+
						(0.020)			**(0.019)**					(0.018)
(BKS) Voluntary withdrawals							**0.056***		0.072					0.059
							(0.027)		(0.048)					(0.047)
(BKS) Failed applications								**0.082***	**0.086***					0.019
								(0.041)	**(0.035)**					(0.040)
(GF) Product suspension prior year										**0.079*** **			**0.057** **	**0.057** **
										(0.017)			**(0.019)**	**(0.019)**
(GF) <183 days since last safety withdrawal											**0.095*** **		**0.074** **	**0.077** **
											(0.024)		**(0.025)**	**(0.025)**
(GF) 183 to 364 days since last safety withdrawal											**0.086*** **		**0.070*** **	**0.070*** **
											(0.022)		**(0.021)**	**(0.020)**
(GF) 365 to 730 days since last safety withdrawal											**0.092*** **		**0.080*** **	**0.080*** **
											(0.020)		**(0.018)**	**(0.017)**
											0.081* **		**0.066*** **	**0.066*** **

Table 19: Coefficients and standard errors of negative binomial regression using standard errors clustered by applicant

	Independent variable: Active review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(GF) >730 days since last safety withdrawal											**(0.019)**		**(0.018)**	**(0.017)**
(GF) Specialization												0.051	0.020	0.018
												(0.034)	(0.028)	(0.028)
Status (ln)	-0.004	-0.005	-0.005	-0.005	-0.004	-0.004	-0.003	-0.003	-0.002	-0.004	-0.002	-0.004	-0.002	-0.002
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
Age (ln)	-0.002	-0.004	-0.003	-0.005	-0.002	-0.003	-0.001	-0.001	0.0001	-0.001	-0.0003	0.002	0.001	0.002
	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Sales (ln)	-0.0004	-0.002	-0.002	-0.003	-0.0005	-0.001	-0.001	-0.001	-0.002	-0.0003	-0.0003	-0.0003	-0.0001	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Applied for AA	**-0.168*** **	**-0.169*** **	**-0.163*** **	**-0.164*** **	**-0.168*** **	**-0.168*** **	**-0.167*** **	**-0.168*** **	**-0.169*** **	**-0.175*** **	**-0.168*** **	**-0.166*** **	**-0.171*** **	**-0.166*** **
	(0.020)	**(0.020)**	**(0.020)**	**(0.020)**	**(0.020)**	**(0.020)**	**(0.020)**	**(0.020)**	**(0.020)**	**(0.019)**	**(0.019)**	**(0.020)**	**(0.019)**	**(0.019)**
Origin	-0.003	-0.005	-0.006	-0.007	-0.003	-0.004	-0.002	-0.005	-0.004	-0.004	-0.002	-0.003	-0.003	-0.004
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.010)	(0.009)	(0.010)
Constant	**5.047*** **	**5.107*** **	**5.068*** **	**5.114*** **	**5.052*** **	**5.066*** **	**5.050*** **	**5.065*** **	**5.077*** **	**5.187*** **	**5.199*** **	**5.010*** **	**5.257*** **	**5.272*** **
	(0.076)	**(0.081)**	**(0.076)**	**(0.081)**	**(0.076)**	**(0.075)**	**(0.075)**	**(0.075)**	**(0.073)**	**(0.088)**	**(0.071)**	**(0.080)**	**(0.078)**	**(0.079)**

*Note: +p<0.1; *p<0.05; **p<0.01; ***p<0.001; (U) = variables measuring product uncertainty, (BK) = variables measuring being known, (BKS) = variables measuring being known for something; (GF) = variables measuring generalized favorability*

Appendix G – Negative binomial regressions with dummies per therapeutic area

Table 20: Regression outcome of negative binomial regression using dummies per therapeutic class

	Independent variable: Active review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
(U) Product complexity (ln)	**0.046*** (0.011)**	**0.038** (0.012)**	**0.043*** (0.011)**	**0.037** (0.012)**	**0.045*** (0.011)**	**0.044*** (0.011)**	**0.044*** (0.011)**	**0.043*** (0.011)**	**0.041*** (0.011)**	**0.028* (0.012)**	0.011 (0.012)	**0.043*** (0.011)**	0.002 (0.013)	-0.0001 (0.013)
(U) Evidence base (ln)	-0.004 (0.004)	-0.002 (0.005)	-0.002 (0.004)	-0.001 (0.005)	-0.004 (0.004)	-0.003 (0.004)	-0.004 (0.004)	-0.004 (0.004)	-0.004 (0.004)	-0.002 (0.004)	0.0001 (0.004)	-0.003 (0.004)	0.001 (0.004)	0.002 (0.004)
(U) <183 days since FDA approval	**-0.025* (0.013)**	**-0.027* (0.013)**	**-0.027* (0.013)**	**-0.028* (0.013)**	**-0.026* (0.013)**	**-0.026* (0.013)**	**-0.026* (0.013)**	**-0.026* (0.013)**	**-0.026* (0.013)**	-0.022+ (0.013)	-0.024+ (0.012)	-0.025+ (0.013)	-0.021+ (0.012)	-0.023+ (0.012)
(U) 183 to 364 days since FDA approval	**0.033* (0.016)**	0.032+ (0.016)	0.031+ (0.016)	0.031+ (0.016)	**0.032* (0.016)**	0.030+ (0.016)	**0.032* (0.016)**	0.031+ (0.016)	0.029+ (0.016)	**0.033* (0.016)**	**0.035* (0.016)**	**0.034* (0.016)**	**0.036* (0.016)**	**0.032* (0.016)**
(U) >365 days since FDA approval	0.009 (0.015)	0.009 (0.015)	0.009 (0.015)	0.009 (0.015)	0.008 (0.015)	0.008 (0.015)	0.008 (0.015)	0.008 (0.015)	0.007 (0.015)	0.013 (0.015)	0.008 (0.015)	0.008 (0.015)	0.011 (0.015)	0.010 (0.015)
(BK) Past approvals (ln)		0.012+ (0.006)		0.009 (0.007)										-0.003 (0.007)
(BK) Current applications (ln)			**0.023* (0.010)**	**0.020* (0.010)**										**0.022* (0.010)**
(BKS) Product referrals					0.021 (0.048)				-0.062 (0.059)					-0.056 (0.057)
(BKS) Involved in safety withdrawal						0.033 (0.021)			0.035 (0.023)					0.034 (0.022)
(BKS) Voluntary withdrawals							0.068 (0.049)		0.086 (0.057)					0.070 (0.054)
(BKS) Failed applications								0.071 (0.049)	0.075 (0.050)					0.025 (0.049)
(GF) Product suspension										**0.044*** (0.012)**			**0.030* (0.013)**	**0.029* (0.013)**
(GF) <183 days since safety withdrawal											**0.088*** (0.017)**		**0.074*** (0.018)**	**0.076*** (0.019)**
(GF) 183 to 364 days since last safety withdrawal											**0.079*** (0.017)**		**0.070*** (0.017)**	**0.068*** (0.017)**
(GF) 365 to 730 days since last safety withdrawal											**0.081*** (0.017)**		**0.074*** (0.018)**	**0.073*** (0.018)**
(GF) >730 days last safety withdrawal											**0.074*** (0.017)**		**0.067*** (0.017)**	**0.067*** (0.017)**
(GF) Specialization												0.032 (0.023)	0.007 (0.023)	0.004 (0.023)
Status (ln)	-0.001 (0.007)	-0.002 (0.007)	-0.002 (0.007)	-0.003 (0.007)	-0.001 (0.007)	-0.001 (0.007)	0.0001 (0.007)	-0.001 (0.007)	0.001 (0.007)	-0.002 (0.007)	-0.0003 (0.007)	-0.001 (0.007)	-0.0004 (0.007)	0.0001 (0.007)
Age (ln)	-0.006	-0.008	-0.007	-0.009	-0.006	-0.007	-0.005	-0.005	-0.003	-0.004	-0.003	-0.003	-0.002	-0.001

	Independent variable: Active review time													
	Controls	1	2	1&2	3	4	5	6	3&4&5&6	7	8	9	7&8&9	All
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Sales (ln)	0.001	-0.001	-0.0003	-0.002	0.001	0.001	0.0004	0.0005	-0.001	0.001	0.001	0.001	0.001	-0.001
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Applied for AA	**-0.159*** **	**-0.160*** **	**-0.153*** **	**-0.154*** **	**-0.159*** **	**-0.159*** **	**-0.158*** **	**-0.160*** **	**-0.160*** **	**-0.164*** **	**-0.161*** **	**-0.158*** **	**-0.164*** **	**-0.158*** **
	(0.019)	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.019)**	**(0.018)**	**(0.018)**	**(0.019)**	**(0.018)**	**(0.018)**
Origin	-0.012	-0.014	-0.015	-0.016	-0.012	-0.012	-0.011	-0.013	-0.012	-0.012	-0.010	-0.011	-0.010	-0.011
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
(A) Alimentary tract	-0.013	-0.010	-0.014	-0.012	-0.013	-0.010	-0.013	-0.012	-0.010	-0.027	0.013	-0.014	0.002	0.003
and metabolism	(0.111)	(0.111)	(0.110)	(0.110)	(0.111)	(0.111)	(0.111)	(0.111)	(0.110)	(0.109)	(0.107)	(0.111)	(0.107)	(0.106)
(B) Blood and blood	-0.019	-0.014	-0.021	-0.017	-0.019	-0.016	-0.019	-0.018	-0.015	-0.037	0.009	-0.020	-0.005	-0.005
forming organisms	(0.111)	(0.111)	(0.110)	(0.110)	(0.111)	(0.111)	(0.111)	(0.111)	(0.110)	(0.109)	(0.108)	(0.111)	(0.107)	(0.106)
(C) Cardiovascular	-0.036	-0.036	-0.035	-0.035	-0.036	-0.034	-0.036	-0.036	-0.033	-0.050	-0.011	-0.035	-0.021	-0.017
system	(0.113)	(0.112)	(0.112)	(0.112)	(0.113)	(0.113)	(0.113)	(0.113)	(0.112)	(0.111)	(0.109)	(0.113)	(0.109)	(0.107)
(D) Dermatologicals	0.005	0.010	-0.001	0.004	0.005	0.008	0.005	0.006	0.009	-0.014	0.029	0.003	0.014	0.011
	(0.115)	(0.115)	(0.115)	(0.114)	(0.115)	(0.115)	(0.115)	(0.115)	(0.114)	(0.114)	(0.111)	(0.115)	(0.111)	(0.110)
(G) Genito urinary	-0.040	-0.035	-0.039	-0.036	-0.040	-0.038	-0.042	-0.040	-0.040	-0.054	-0.017	-0.042	-0.028	-0.028
system	(0.113)	(0.112)	(0.112)	(0.112)	(0.113)	(0.113)	(0.113)	(0.113)	(0.112)	(0.111)	(0.109)	(0.113)	(0.109)	(0.108)
(H) Hormonal	-0.024	-0.027	-0.028	-0.029	-0.026	-0.022	-0.028	-0.024	-0.020	-0.032	0.003	-0.026	-0.004	-0.003
preparations	(0.114)	(0.114)	(0.114)	(0.113)	(0.114)	(0.114)	(0.114)	(0.114)	(0.114)	(0.113)	(0.110)	(0.114)	(0.110)	(0.109)
(J) Antiinfectives	-0.092	-0.088	-0.095	-0.092	-0.092	-0.089	-0.094	-0.090	-0.088	-0.097	-0.055	-0.090	-0.060	-0.062
	(0.111)	(0.111)	(0.111)	(0.111)	(0.111)	(0.111)	(0.111)	(0.111)	(0.111)	(0.110)	(0.108)	(0.111)	(0.107)	(0.106)
(L) Antineoplastic	-0.048	-0.047	-0.050	-0.049	-0.049	-0.047	-0.051	-0.050	-0.049	-0.063	-0.022	-0.048	-0.032	-0.034
	(0.111)	(0.110)	(0.110)	(0.110)	(0.111)	(0.110)	(0.110)	(0.110)	(0.110)	(0.109)	(0.107)	(0.110)	(0.106)	(0.105)
(M) Musculoskeletal	-0.023	-0.015	-0.033	-0.025	-0.024	-0.023	-0.025	-0.025	-0.025	-0.042	-0.003	-0.026	-0.017	-0.029
system	(0.115)	(0.115)	(0.115)	(0.114)	(0.115)	(0.115)	(0.115)	(0.115)	(0.114)	(0.114)	(0.111)	(0.115)	(0.111)	(0.110)
(N) Nervous system	-0.014	-0.009	-0.014	-0.010	-0.015	-0.013	-0.015	-0.016	-0.014	-0.031	0.003	-0.017	-0.010	-0.009
	(0.112)	(0.112)	(0.111)	(0.111)	(0.112)	(0.112)	(0.112)	(0.112)	(0.111)	(0.110)	(0.108)	(0.112)	(0.108)	(0.107)
(P) Antiparasitic	0.014	0.031	0.021	0.033	0.014	0.018	0.014	0.018	0.021	-0.042	0.019	0.008	-0.014	-0.009
products	(0.153)	(0.152)	(0.152)	(0.152)	(0.153)	(0.152)	(0.152)	(0.152)	(0.152)	(0.151)	(0.148)	(0.152)	(0.147)	(0.146)
(R) Respiratory	0.011	0.006	0.006	0.003	0.011	0.008	0.009	0.013	0.009	-0.018	0.032	0.009	0.012	0.004
system	(0.115)	(0.115)	(0.115)	(0.114)	(0.115)	(0.115)	(0.115)	(0.115)	(0.115)	(0.114)	(0.112)	(0.115)	(0.111)	(0.110)
(S) Sensory organs	-0.035	-0.033	-0.035	-0.034	-0.034	-0.033	-0.035	-0.035	-0.035	-0.060	-0.019	-0.035	-0.036	-0.037
	(0.113)	(0.113)	(0.112)	(0.112)	(0.113)	(0.113)	(0.113)	(0.113)	(0.112)	(0.112)	(0.109)	(0.113)	(0.109)	(0.108)
(V) Various	-0.037	-0.037	-0.038	-0.038	-0.037	-0.035	-0.038	-0.036	-0.034	-0.049	-0.019	-0.034	-0.027	-0.027
	(0.112)	(0.112)	(0.112)	(0.111)	(0.112)	(0.112)	(0.112)	(0.112)	(0.112)	(0.111)	(0.109)	(0.112)	(0.108)	(0.107)
Constant	**5.047*** **	**5.103*** **	**5.071*** **	**5.111*** **	**5.052*** **	**5.060*** **	**5.051*** **	**5.062*** **	**5.068*** **	**5.154*** **	**5.173*** **	**5.024*** **	**5.222*** **	**5.240*** **
	(0.128)	**(0.131)**	**(0.128)**	**(0.130)**	**(0.128)**	**(0.128)**	**(0.128)**	**(0.128)**	**(0.128)**	**(0.130)**	**(0.126)**	**(0.129)**	**(0.129)**	**(0.130)**
Observations	488	488	488	488	488	488	488	488	488	488	488	488	488	488
R^2	0.236	0.241	0.243	0.245	0.237	0.239	0.239	0.240	0.245	0.259	0.284	0.240	0.294	0.304
Log Likelihood	-219.955	-2189.193	-2188.207	-2187.215	-2190.855	-2189.765	-2190.002	-2189.910	-2187.654	-2184.532	-2175.395	-2190.012	-2172.567	-2167.895
AIC	4,431.909	4,430.387	4,428.414	4,428.430	4,433.711	4,431.531	4,432.004	4,431.821	4,433.308	4,421.064	4,408.790	4,432.024	4,407.135	4,409.790

*Note: +p<0.1; *p<0.05; **p<0.01; ***p<0.001; (U) = variables measuring product uncertainty, (BK) = variables measuring being known, (BKS) = variables measuring being known for something; (GF) = variables measuring generalized favorability*