Potential to Simplify the Writing of Submission Documents: Evaluation of Publicly Available Module 2 Documents in Drug Submissions from Different Pharmaceutical Companies

Master's thesis by

Ronja Alexandra Forstén

Pharmaceutical Sciences
Faculty of Science and Engineering
Åbo Akademi University
Åbo, Finland
2021

Acknowledgements

First and foremost, I am extremely grateful to Bayer for giving me the golden opportunity to proceed with this research and for the financial support. I would also like to express my special thanks to my supervisor Päivi Norja at Bayer - the completion of the research project would not have been possible without your incomprehensible support, guidance, encouragement, and willingness to help.

I would like to extend my sincere thanks to Heini Metso, Walther Seiler and Allison Kelly who took their time to contribute with valuable insights and senior expertise.

I would also like to thank Marion Hardtke and Tuomas Heikkinen for their expert advice and wonderful collaboration throughout the research.

Many thanks to my professors Jessica Rosenholm and Outi Salo-Ahen at Åbo Akademi University for the support and flexibility throughout this research project.

Finally, I would like to express my gratitude to each and every person who somehow contributed to this research project.

Abstract

MSc Degree: Drug Development and Medical Technology	
Author: Ronja Alexandra Forstén	
Thesis title: Potential to Simplify the Writing of Submission Documents: Evaluation of Publicly Available Module 2 Documents in Drug Submissions from Different Pharmaceutical Companies	
Supervisors: Jessica Rosenholm, Outi Salo-Ahen (Åbo Akademi University)	Supervisors: Päivi Norja, Heini Metso, Marion Hardtke (Bayer Pharmaceuticals)

Documentation of clinical trials is surrounded by a jungle of directives and recommendations, which are constantly growing. It has shown to be troublesome to interpret what content the regulatory authorities want in the different submission documents. Challenging interpretation of regulations and directives has left greater writing differences between sponsors and it is important to implement new writing strategies that take these regulatory changes into consideration. The writing strategy must focus on the content necessary to grant a market approval as well as content demonstrating the benefit of the drug. Additionally, transparency has become an essential part of the drug submission process.

The purpose of this research project was to investigate how different sponsors write their publicly available clinical module 2.5 and 2.7 documents, and how well the CTD (Common Technical Document) guideline was followed. Moreover, simplification suggestions were prepared for the writing of module 2 documents based on the result findings.

Ten sponsors and 20 module 2 packages were randomly selected according to the scope of the thesis. One package was downloaded from EMA's (European Medicines Agency) website and one package from Health Canada's website per sponsor. The study design was divided into a literature review followed by a quantitative evaluation and comparison of different variables, including number of pages, tables and figures, number and structure of level-1 headings and subheadings, and categorization of table content repetition in the document text.

The results demonstrated that all CTD recommendations were generally followed, with an exception of the recommendation for the document length. Various industry-wide simplification suggestions were gathered from the results. The main suggestion was to reduce the document length for both module 2.5 and 2.7 documents, in accordance with the CTD guideline. To reduce the document length, it would be beneficial to consider to only generate essential tables and figures, and to reflect what has been presented in the table content compared to the body text. Other recommendations would be to reduce repetition, to include cross-referencing and to consider whether it is vital to include extra subheadings that are not recommended by CTD.

Simplifying the document writing process could impact resources and document quality and increase trust between the public and the sponsor. All simplification suggestions might not apply to every sponsor, although it is obvious that all sponsors should examine their approaches and consider whether their documents could be simplified to a certain degree.

Date: 7.5.2021	Pages: 114

Table of Contents

Table of Tables

Table of Figures

List of abbreviations

Abbreviation	Full description
CCI/CBI	Commercially Confidential Information (EMA)/Confidential Business Information (Health Canada)
CDP	Clinical Data Publication
CHMP	Committee for Medicinal Products for Human Use
CSR	Clinical Study Report
CTD	Common Technical Document
ECRIN	European Clinical Research Infrastructures Network
EEA	European Economic Area
EMA	European Medicines Agency
EU	the European Union
EU	European Union
FDA	Food and Drug Administration
FIH	First in human
HC	Health Canada
HC PRCI	Health Canada Public Release of Clinical Information
ICH	the International Conference on Harmonisation
IFPMA	International Federation of Pharmaceutical Manufacturers Associations
IPD	Individual Patient Data (EMA)

JPMA	Japan Pharmaceutical Manufacturers Association
MA	Marketing Authorization
MAA	Marketing authorization Application
MHLW	Ministry of Health, Labor, and Welfare (Japan)
NDS	New Drug Submission
NDS-NAS	New Drug Submission - New Active Substance
PhRMA	Pharmaceutical Research and Manufacturers of America
PK	Pharmacokinetic
PMDA	Pharmaceuticals and Medical Devices Agency
PPD/PI	Protected Personal Data (EMA)/Personal Information (Health Canada)
SAE	Serious Adverse Event
the US	the United States of America
WHO	the World Health Organization

1 Introduction

In the pharmaceutical industry, good documentation practice is a requirement when handling and recording data of a drug product to ensure quality. The simple rule is "if it isn't documented, it didn't happen" (Vuolo, 2020), but could it be possible that excessive resources are put into writing a document?

The number of regulations and recommendations are constantly growing. Increased interpretation when writing the documents has become a challenge that should be addressed. For some sponsors[1] (meaning a pharmaceutical company) interpretation leaves room for inclusion of redundant information, i.e., over-explanation and unnecessary repetition, in the documents of the MAA (marketing authorization application).

When a sponsor is applying for a marketing authorization of a drug product[2] in Europe, i.e., EMA (European Medicines Agency), an MAA must be submitted. An MAA contains many different documents from both pre-clinical and clinical studies and all documentation necessary for an MAA is structured in five different modules (1 to 5), according to the CTD (Common Technical Document) guidelines:

- Module 1 - region-specific administrative and prescribing information
- Module 2 - overviews and summaries of module 3 to 5
- Module 3 - quality of pharmaceutical documentation
- Module 4 - non-clinical reports of pharmacology and toxicology
- Module 5 - clinical study reports from clinical trials. (EMA, 2004).

The regulatory authority will review the data from these studies and a recommendation for a marketing authorization is based on a voting system, which will disclose a positive or negative opinion. The sponsor must submit a proposal package of redacted and anonymized versions of the clinical documents from the submitted application. This package is to be made publicly available soon after the authority decision on the marketing authorization or withdrawal of the application by the sponsor themselves. The publicly

[1] A sponsor is the leading company that is financially supporting an activity or event, often a pharmaceutical company

[2] The EMA term for drug product is "medicinal product", however, the general term "drug" is used throughout the thesis

available documents are the module 2 and 5 clinical documents from the CTD structure. The availability is the result of what is known as transparency rules. (EMA, 2004; EMA, 2018).

Transparency has taken up a major role in the pharmaceutical industry and the reasons for this are many. According to the World Health Organization, a key element to create safe, effective and good-quality drugs is transparency in governance. Transparency is also an inevitable prerequisite of fair pricing (WHO, 2018). In this respect, transparency is vital to create trust between the public and the pharmaceutical industry and to provide opportunities for anyone to analyze different studies and to form a personal perception of a particular product. Transparency will also force innovative thinking and create research opportunities for other companies. The transparency rules are continuously being updated and implemented to an increased degree, with the latest rules for Europe and Canada being:

- EMA Clinical Data Publication, also known as Policy 0070
- HC (Health Canada) PRCI (Public Release of Clinical Information) (EMA, 2019; Health Canada, 2019a).

The guidelines prepare the sponsor for publishing module 2 and 5 documentation from the MAA. In other words, these documents are publicly available to be analyzed and downloaded from portals maintained by EMA and HC, and it is important to keep in mind that the information in these public documents can be read by anyone. (EMA, 2019; Health Canada, 2019a).

Before publishing the documents, personal data and confidential information must be redacted or anonymized to protect the individuals who participated in or conducted the clinical trials that the MAA is based on. Protection of the patients who participated in the trials is paramount. In addition, other information (such as previously mentioned unnecessary repetition) that is considered redundant to grant a marketing authorization could be removed. (EMA, 2019; Health Canada, 2019a).

For this research, publicly available module 2 clinical documents were chosen for analysis. Module 2 is divided into seven subsections, however, only module 2.5 and 2.7

are made publicly available. Module 2.5 presents a clinical overview whilst module 2.7 is split into four different documents: 2.7.1 Summary of Biopharmaceutic Studies and Analytical Methods, 2.7.2 Summary of Clinical Pharmacology Studies, 2.7.3 Summary of Clinical Efficacy and 2.7.4 Summary of Clinical Safety.

Altogether 20 module 2 packages from the MAA, published by EMA or HC and prepared by 10 different pharma sponsors, were chosen for the thesis evaluation. One module 2 package refers to the publicly available documents of the module sections 2.5, 2.7.1, 2.7.2, 2.7.3, and 2.7.4. All EMA packages were initial MAA packages and the HC packages were either NDS (New Drug Submission) or NDS-NAS (New Drug Submission - New Active Substance). Selected variables were compared between the documents and, with this approach, the writing strategies of different sponsors were investigated.

The following objectives of the research were studied:

- To measure structural differences in 20 module 2 packages from 10 different sponsors
- To measure data repetition between tables and paragraph text in module 2.5 and 2.7.4 documents
- To evaluate suggestions for possible documentation simplification

The aim of this research project has been to consider whether the writing strategies could be simplified and if any simplification suggestions could be implemented for the preparation of publicly available module 2 documents. The documents were compared, their relation to the sponsors was analyzed, and the results were compared to the CTD guidelines for module 2 documents to investigate how well the guideline was followed.

This is a general analysis of the industry to create a baseline for future awareness and process changes. It is essential to understand the importance of strategies being implemented in early writing and to realize the necessity of not only writing a good document, but also understanding how certain information is handled and what kind of processes take place after the document has been written.

The research is built on both old and relatively new regulations and guidelines. The author guides the reader through the drug development phases until marketing authorization and includes the appropriate rules and regulations as the timeline is explained (Figure 1).

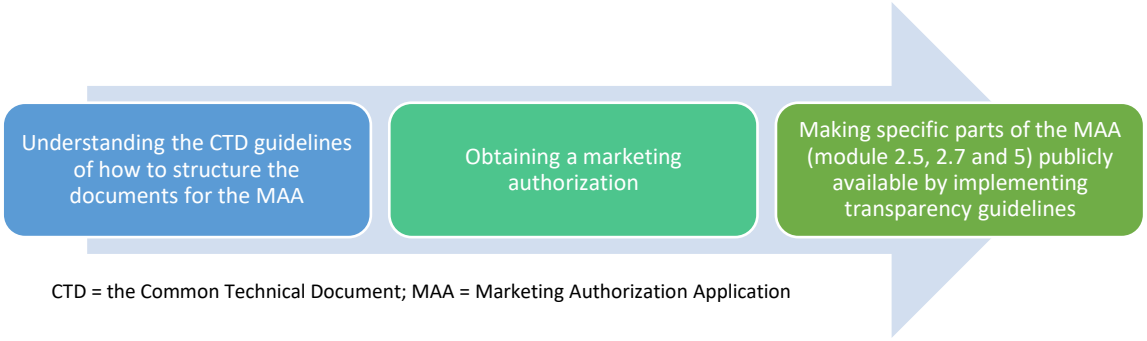

Figure 1. The above-mentioned themes are discussed in the literature review.

As this research is based on new research, the reliability of the results and conclusions are formed on the author's own results along with forthcoming discoveries from future articles, leaving room for development.

The research is a collaboration with the Medical Writing function within Bayer Pharmaceuticals. During the research period, the author has been working as an Associate Medical Writing Specialist alongside the academic work.

The author is writing from a European perspective, unless mentioned otherwise.

1.1 Background

A drug's lifecycle (Figure 2) starts in the research and development (R&D) department. When an active ingredient has been identified and combined with excipients, a drug is created. The optimized drug is then evaluated through various pre-clinical testing stages:

- *In silico* - computerized testing
- *In vitro* - tests performed outside of a living organism (e.g., tubes, glasses, agar plate)
- *In vivo* - tests performed in living organisms (animals)

The data are gathered from pre-clinical testing and sent to the authorities and if the results are promising enough, a permission to start clinical trials is granted. (E.L. Andrade, 2016).

Clinical trials are conducted in humans to investigate the best way of diagnosing and treating a disease. Scientists and doctors are operating the trials, which are divided into four phases described in Table 1. Note that the number of humans in each phase is an approximation per study. However, one complete phase 1 package will include numerous studies. (David J. Kerr, 2006).

Table 1. A brief overview of the different phases of a clinical trial (David J. Kerr, 2006)

Trial phase	**Pre-clinical phase**	**Phase 1**	**Phase 2**	**Phase 3**	**Phase 4**
Number of humans	• No humans involved • In silico, *in vitro* and animal *in vivo* studies	• 10-80 humans • Healthy humans • Humans with given condition if HIV drug, cancer drug, life threatening disease or drugs with SAEs	• 50-200 humans • Humans with given condition	• several hundred - several thousand humans • Humans with given condition	• Hundredths of thousands of humans • Humans with given condition
Purpose	• Toxicology, animal pharmacology and safety studies • To determine a safe FIH dose	• Focus on safety, drug dose, metabolism studies (PK studies and metabolites)	• Explore efficacy and safety	• Confirmatory studies of safety and efficacy • Comparison of existing drug (if any) with new drug	• Continuous drug monitoring after MA has been granted • To bring attention to doctors, detect unexpected adverse events and look for new indications

FIH = First in Human; HIV = Human Immunodeficiency Virus; PK = Pharmacokinetic; SAE = Serious Adverse Event; MA = Marketing Authorization

When clinical trials are carried out, extensive work is required in collecting samples and data, composing documents, and following guidelines. All trials are recorded and documented in a detailed and regulated way. When all the necessary information is assembled in a dossier and guidelines applied, a submission-ready marketing authorization application package is compiled, hereinafter referred to as an MAA. This package, which may include thousands of pages of text, numerous graphs, figures, and other important information, will be submitted to and reviewed by the authority concerned, and if the

collected information is qualified, a MAA can be granted. (EMA, 2019; Health Canada, 2019a).

Certain documents from the MAA are collected in databases that allow the public to view the documentation. The publicly available clinical documents (module 2.5, 2.7 and module 5 CSRs (clinical study reports)) contain sensitive information and must comply with transparency guidelines, which involve guidance on redaction and anonymization approaches before publishing. (EMA, 2019; Health Canada, 2019a).

Figure 2. Process flow chart from drug discovery to marketing authorization.

2 Literature review

This literature review is a selective approach, where a limited number of guidelines from Europe and Canada are analyzed, related to the CTD (Common Technical Document), MAA and new transparency regulations.

The following documents have been studied:

- ICH Topic M4, Common Technical Document for the Registration of Pharmaceuticals for Human Use – Organization Common Technical Document. (CPMP/ICH/2887/99)
- Obtaining an EU marketing authorisation, step-by-step by EMA
- European Medicines Agency policy on publication of clinical data for medicinal products for human use - EMA/144064/2019
- External guidance on the implementation of the European Medicines Agency policy on the publication of clinical data for medicinal products for human use - EMA/90915/2016 (version 1.4)
- Public Release of Clinical Information: guidance document by Health Canada (version 1.0)

2.1 Common Technical Document (CTD)

The data from pre-clinical and clinical phases are presented in various document types and these are filed to the regulators as part of the MAA. All documents submitted to the authority must follow a certain structure and these structural guidelines are often provided by ICH (International conference on Harmonisation), although some countries follow their own national guidelines. (ICH, n.d.).

ICH has provided over 60 different guidelines for the pharmaceutical industry. These guidelines include quality, safety, efficacy, multidisciplinary (cross-cutting topics), MedDRA[3] and CTD. All guidelines are equally important, although the emphasis in this thesis is on CTD and its common format to accumulate all the quality, safety and efficacy data. The European Regulatory Authorities require all MAA documentation to be CTD ready (EMA, 2004). The regulatory authorities that are following the ICH guidelines operate with the same document format (ICH, n.d.).

ICH is responsible for the maintenance of the different guidelines, including the guideline for the CTD dossier. Europe, Japan and the US all have their specific directions for what submission-ready documents should look like. This involves, for example, directions for summaries and table preparations. The CTD was designed by implementing all three region-specific regulations to ensure that the structure of documents is comparable. Today, the CTD format has also been adopted in other countries, such as Canada and Switzerland. (Jordan, 2014).

The CTD guideline contains specifications for how sponsors are recommended to create documents for an MAA. The CTD dossier comprises five different modules. These modules all represent a certain section necessary for registration of human pharmaceuticals. Every module contains detailed instructions on how to format the documents. When companies follow the CTD standards, time and resources are reduced, as all documents have the same formatting structure. A global standard makes it easier to read and find information, both for the regulatory authorities and among companies. (EMA, 2004, p. 3).

[3] MedDRA is a standardized medical terminology dictionary

The general rules concerning the CTD are available and are as follows:

- Style and size of the font should be easily readable and must therefore be large enough, particularly text within tables or figures
- Recommended style in narrative text is Times New Roman with a 12-point font
- Document information must be clear, transparent and easy to review
- All pages must follow sequential numbering, and acronyms as well as abbreviations are to be spelled out the first time when mentioned in a module
- Margins are to be arranged with the aim of printable documentation in formats of A4 paper in Europe and Japan; 8.5 x 11" paper in the US
- To avoid text being obscured in documents that are banded together, the left margin should be adequately wide
- References are cited according to the latest version of ICMJE (Uniform Requirements for Manuscripts Submitted to Biomedical Journals, International Committee of Medical Journal Editors). (EMA, 2004, pp. 3-4).

The five different sections of the CTD dossier are module 1 to 5 (Figure 3): 1 - Administrative information and prescribing information, however, this module is region specific and the content and format do not follow CTD specific rules; 2 - overviews and summaries of module 3 to 5; 3 - quality of pharmaceutical documentation; 4 - non-clinical reports of pharmacology and toxicology; 5 - clinical study reports from clinical trials. (EMA, 2004, p. 4).

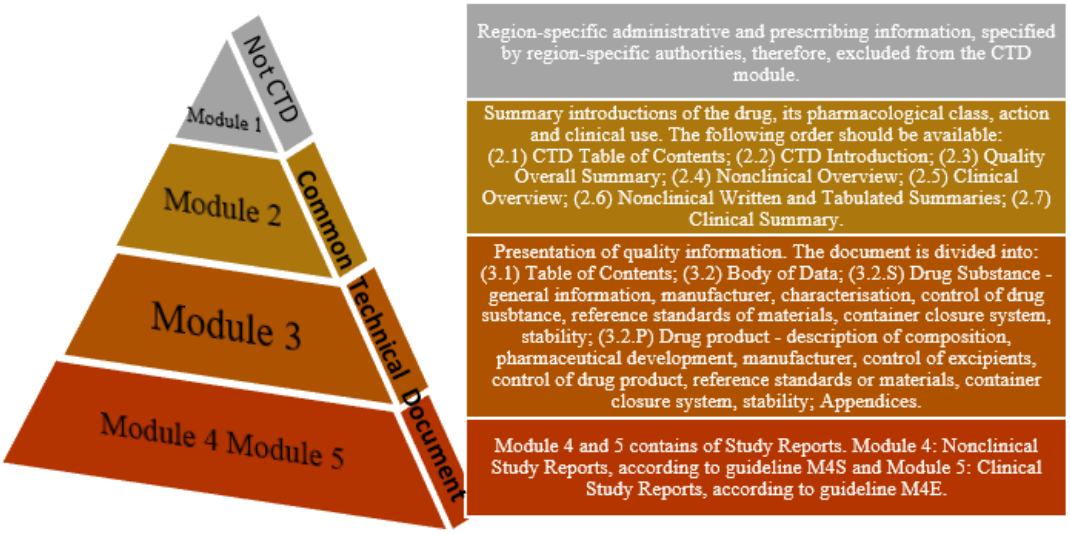

Figure 3. The Common Technical document can be illustrated in a triangle figure. The figure is a re-creation as a summary of ICH M4Q, M4S and M4E guidelines. This thesis focuses on module 2 documents.

All modules are further divided into sections and subsections. However, as the scope of the thesis is publicly available module 2 documents, the focus is on the structure of these documents. The documents are summary introductions of the drug, its pharmacological class, action and clinical use. The following content is available in module 2: (2.1) CTD Table of Contents, (2.2) CTD Introduction, (2.3) Quality Overall Summary, (2.4) Nonclinical Overview, (2.5) Clinical Overview, (2.6) Nonclinical Written and Tabulated Summaries, (2.7) Clinical Summary. Additionally, each and every one of these documents contains level-1 headings (see section 4.3) and sponsors might also choose to include additional level-1 headings, which are out of scope for the CTD structure. These level-1 headings may also be further divided into subheadings, depending on the document. (EMA, 2016).

Only two of the seven documents in module 2 are publicly available: Clinical Overview (2.5) and Clinical Summary (2.7) documents. Their sections and level-1 headings are presented in Table 2. Note that the clinical overview only contains one document, whilst the clinical summaries are divided into four separate documents. (EMA, 2016).

In accordance with the CTD recommendations, the clinical overview should be a short but detailed ca. 30-page description of the development plan, which involves safety, efficacy

risk analyses, biopharmaceutics and clinical pharmacology. The development plan of evidence shows completed, initiated and planned studies. The clinical summaries are more detailed than the clinical overview and document length of the full module 2.7 summary document (module 2.7.1, 2.7.2, 2.7.3, 2.7.4 combined) is recommended to be between 50 and 400 pages. (EMA, 2016).

Table 2. Publicly available module 2 documents and their level-1 headings (EMA, 2016)

2.5 Clinical Overview	**2.7.1 Summary of Biopharmaceutic Studies and Analytical Methods**	**2.7.2 Summary of Clinical Pharmacology Studies**	**2.7.3 Summary of Clinical Efficacy**	**2.7.4 Summary of Clinical Safety**	**Conclusion**
2.5.1 Product development rationale	2.7.1.1 Background and overview	2.7.2.1 Background and overview	2.7.3.1 Background and overview of clinical efficacy	2.7.4.1 Exposure to the drug	Nothing is specified about including a specific level-1 conclusion heading
2.5.2 Overview of biopharmaceuticals	2.7.1.2 Summary of results of individual studies	2.7.2.2 Summary of results of individual studies	2.7.3.2 Summary of results of individual studies	2.7.4.2 Adverse events	**Reference list**
2.5.3 Overview of clinical pharmacology	2.7.1.3 Comparison and analyses of results across studies	2.7.2.3 Comparison and analyses of results across studies	2.7.3.3 Comparison and analyses of results across studies	2.7.4.3 Clinical laboratory evaluations	One reference list must be included in module 2.5 and at least one must be provided for all summary documents (module 2.7)
2.5.4 Overview of efficacy		2.7.2.4 Special studies	2.7.3.4 Analysis of clinical information relevant to dosing recommendations	2.7.4.4 Vital signs, physical findings, and other observations related to safety	**Appendix**
2.5.5 Overview of safety			2.7.3.5 Persistence of efficacy and/or tolerance effects	2.7.4.5 Safety in special groups and situations	Included in the end of a document if detailed presentation of methods and results. If tables are lengthy they should be provided in the appendix
2.5.6 Benefits and risks conclusions				2.7.4.6 Post-marketing data	

2.2 Obtaining marketing authorization in Europe

MAA evaluation is the process that a new drug must go through in order to have a permission to place the drug on the market. A marketing authorization in the EU may result from a centralized procedure, a decentralized procedure, mutual recognition procedure or a national authorization procedure. (EMA, 2019). A more detailed description is found in Table 3.

Table 3. Short description of the different MAA procedures

Centralized procedure	Decentralized procedure	Mutual recognition procedure	National authorization procedure
• A single MAA to EMA, granted in the whole EU + EEA • This is compulsory for certain medicines (cancer, diabetes, HIV, viral diseases, neurodegenerative diseases, auto-immune and other immune dysfunctions) • The benefit of this procedure is centralized safety monitoring and product information in all EU languages **(EMA, 2019)**	• No MA has been granted in a member state before and cannot be granted through the centralized procedure • This can be applied in several member states at the same time (EMA, 2007)	• MA has previously been granted in a member state and by recognition can be granted in another state (EMA, 2007)	• Each member state has its own national authorization • This can be used if a sponsor is intending to market the drug in only one of the member states • The application must be submitted in the state's own language (EMA, 2019)

MA = Marketing Authorization; MAA = Marketing Authorization Application; EU = European Union; EMA = European Medicines Agency; EEA = European Economic Area; HIV = Human Immunodeficiency Virus

As most of today's new medicines are granted marketing authorization via the centralized procedure, the focus will be on this route. In addition, EMA Policy 0070 only applies to the centralized procedure applications. (EMA, 2019).

The MAA can moreover be divided into three different approval types: standard, conditional approval or exceptional circumstances (see Table 4). (EMA, 2015).

Table 4. Types of marketing approval (EMA, 2015)

Approval type	Standard	Conditional approval	Exceptional circumstances
Clarification	• Normal approval where comprehensive efficacy, safety and quality data have been justified	• Comprehensive data are not available yet • Granted if unmet need must be fulfilled • Risk benefit profile approved • Approval valid for 1 year with possibility for renewal	• Comprehensive data are not available and cannot be made available (for efficacy and safety) • Rare indication • Ethical barriers to collect data • Certain criteria must be met

This thesis concentrates mainly on the European authority processes, but it is good to remember that various authorities process the MAAs depending on the region in which the approval is applied. In Europe, the applications are handled by EMA, but in Canada similar applications are handled by Health Canada and in the US by FDA[4]. When a pharmaceutical company applies for an authorization to market its drug, the applicant must consider in which region to hold the authorization and which specific rules are to be followed. The regulatory authority will review and possibly approve the request once the applicant has a drug MAA prepared. For all authorities, a submission arrangement consists of a pre-submission meeting, administrative review, agency review, sponsor response and agency decision. (EMA, n.d. (c); Health Canada, 2019b; FDA, New Drug Application (NDA), 2019).

2.2.1 Marketing authorization process

When phase 1 to 3 trials are conducted and a favorable benefit-risk profile[5] is proven, a product is ready to be marketed (Patrick Waller, 2017). The market evaluation starts with submission of CMC quality data (meaning chemistry, manufacturing and controls data to ensure safety, efficacy and batch consistency) and data from pre-clinical and clinical phases. The data are evaluated by the CHMP (Committee for Medicinal Products for Human Use), i.e., a function within EMA. Additional experts might be consulted by EMA as needed during this process and the sponsor can be asked to complement the already submitted data with clarifications or additional analysis. At the end of the process, a

[4] NDA (New Drug Application) is the FDA corresponding term to EMA's MAA

[5] A benefit-risk profile is a way to measure the safety of a drug by weighing the risks of a drug against the benefits

decision is made by a formal vote within the CHMP, and the product is either approved or refused. However, as EMA has no power to make a final decision under EU law, the recommended decision is forwarded to the European Commission which then makes a final legally binding decision (Figure 4) (EMA, 2019).

Figure 4. Marketing authorization process flowchart (EMA, 2019).

2.2.2 Submission flowchart

As seen in Figure 5, the centralized procedure for marketing authorization is a review process lasting 201 days, excluding the time it takes for the sponsor to provide feedback. The sponsor sends an MAA to EMA on day 0 and within two weeks EMA should validate the application. If the application meets all necessary validation criteria it is forwarded to the CHMP, which starts the evaluation process. The day the CHMP initiates the data evaluation, the review clock starts from day 0. The committee has 120 days to evaluate the data and questions are then send back to the sponsor. When the sponsor receives the feedback, the review clock stops, and it is only continued once the responses are provided by the sponsor back to EMA. The sponsor has a maximum of three months to submit the answers. When the sponsor has forwarded the answers to EMA, the responses are handled by the CHMP and the review clock continues (day 121). The CHMP then has 89 days to come to a final decision - approval or refusal (day 210). (EMA, 2019).

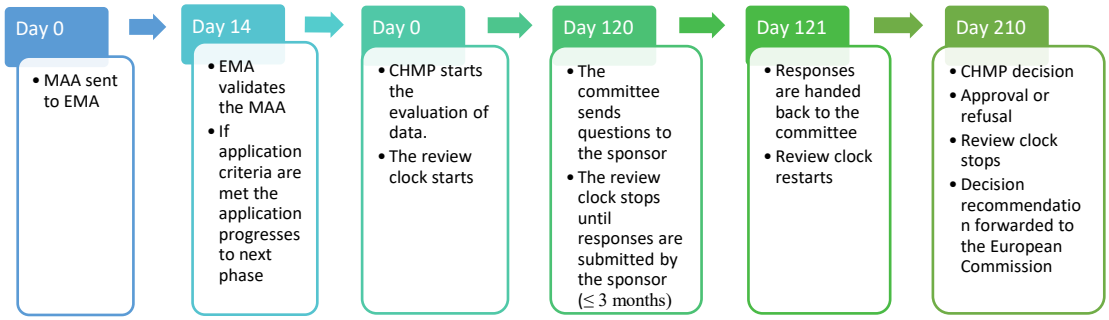

Figure 5. Centralized procedure of marketing authorization.

If the CHMP decides to refuse the marketing authorization, the sponsor has the right to appeal and garner a second opinion. When and if a drug is approved by the CHMP, essential information is passed on to the European Commission, the member states and the applicant[6]:

- CHMP's opinion
- An assessment report
- A summary of product characteristics
- Labelling
- Package. (EMA, 2019).

After the information transfer, EMA will publish an EPAR (European Public Assessment Report). This report is published for drugs that are granted a MA through the centralized procedure (also used for suspended or withdrawn products). In this report, a simple process overview summary (in a question and answer format) and package leaflet are found. (EMA, 2015).

2.2.3 Involved committees in the MAA process

CHMP is involved in the regulatory process together with multiple other committees, as applicable. COMP (Committee for Orphan Medicinal Products), CAT (Committee for Advanced Therapies) and PDCO (Paediatric Committee) are committees that can be involved in pre-submission processes. CHMP, CAT, PDCO, PRAC (Pharmacovigilance Risk Assessment Committee) and COMP can participate in the MAA evaluation phase and CHMP and PRAC might also be included in later post marketing authorization processes. See Figure 6. (EMA, 2015; EMA, n.d. (a)).

[6] The applicant is the organization or legal person who has submitted the clinical reports to the authority/authorities

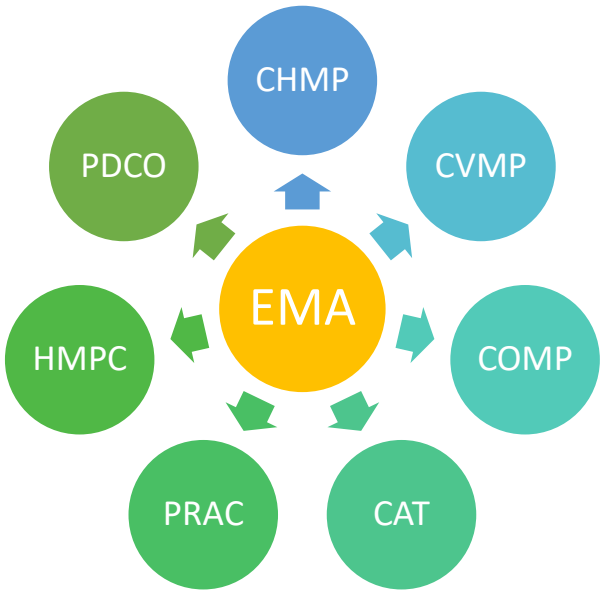

Figure 6. EMA has different competent committees handling specific matters (EMA, 2015).

HMPC = Committee for Herbal Medicinal Products; PDCO = Paediatric Committee; CHMP = Committee for Human Medicinal Products; CVMP = Committee for Veterinary Medicinal Products; COMP = Committee for Orphan Medicinal Products; CAT = Committee for Advanced Therapies; PRAC = Pharmacovigilance Risk Assessment Committee

2.3 Transparency in Europe and Canada

Today, it is truly important that clinical information, to a great extent, reaches the public domain and that region-specific transparency guidelines are followed before information is published. The reasons for transparency are mainly:

- To create trust between the public and the pharmaceutical industry
- To provide opportunities for anyone to analyze different studies and to create a personal perception of a particular product
- To force innovative thinking and create research opportunities for other companies. (European Clinical Research Infrastructure (ECRIN), n.d.; EMA, n.d. (d)).

Innovative thinking and strategy changes are key for pharma companies when applying the transparency policies. Previously, the submission-ready documents were only seen by the authorities and were closed for the general public, which lead to situations where excessive information might have been put into the documents, for different reasons. When applying the transparency guidelines, the strategy has had to change, and only vital information is to be put into the documentation. This way fewer resources are required for retrospective work. (personal communication, 31 Jul 2020).

Europe and Canada have determined to make drug approval decisions as accessible as possible, and data collection and processing has formed what the transparency guidelines are today. (Health Canada, 2019a; EMA, n.d. (d)).

The transparency guidelines concern documents that are to be published after an MAA has been approved, refused or withdrawn. The publicly available clinical documents are found on EMA's and Health Canada's websites, respectively. The authorities handle the regulations regarding the transparency of the documents. Both authorities have their specific rules and the sponsor must ensure that these rules are properly implemented for each region. Along with document publishing, the sponsor must ensure that sensitive information is anonymized or redacted. (Health Canada, 2019a; EMA, n.d. (d)).

A various number of guidelines regarding transparency has been written and applied throughout the years. Continuous updates and new guidelines have led to where we stand today. The requirements and the guidelines will continuously develop as the authorities gain further experience and have a more complete documentation systems for publishing. The final result of this implementation is expected to bring about publications of all clinical trials, regardless of the results. (European Clinical Research Infrastructure (ECRIN), n.d.; EMA, n.d. (d)).

The development of clinical trial transparency can be divided into three phases Figure 7:

Figure 7. Implemented guidelines in Europe throughout the years. These are only examples of many implementations (NIH; EMA, 2019; EudraCT, n.d.).

EudraCT stands for European Union Drug Regulating Authorities Clinical Trials and is a database where data are collected on drugs in clinical trials in EU and EEA (EudraCT, n.d.). The clinical trials register holds information on clinical trials conducted in the EU and EEA, these results are then entered to the EudraCT database (EMA, 2021a). The latest guidelines for transparency have been implemented in Europe (EMA Policy 0070) as recently as 2015. The implementation of these guidelines has made it possible for this research to exist, as module 2.5 and 2.7 documents now must be publicly available and, therefore, can be analyzed (European Clinical Research Infrastructure (ECRIN), n.d.; EMA, n.d. (d)).

2.4 EMA Policy 0070

EMA Policy 0070, also known as Clinical Data Publication (CDP), refers to the European Medicines Agency and their policy on the publication of clinical data for medicinal products for human use. EMA Policy 0070 is the flagship policy and reason for the opportunity to download module 2 documents from the clinical data publication portal of EMA.

2.4.1 Brief introduction

The policy is divided into two phases: phase 1 and phase 2. Phase 1 has been implemented since January 2015, and phase 2 will only come into effect later. Hence, in scope of this thesis is information associated with phase 1 alone.

EMA Policy 0070 and the external guidance documents are available to provide assistance on how to manage anonymization and redaction processes before the documents are to be published on EMA's website. The policy provides guidance on anonymization of PPD (Protected Personal Data), identification and redaction of CCI (Commercially Confidential Information) in clinical reports, and practical standpoints when submitting clinical reports. These are all necessary processes that must be accomplished before the documents can be published in the EMA portal. (EMA, 2018, p. 6; EMA, 2019).

An external guidance document "*External guidance on the implementation of the European Medicines Agency policy on the publication of clinical data for medicinal products for human use*" was created to further clarify how to implement EMA Policy 0070, which includes following information:

- External guidance on the procedural aspects related to the submission of clinical reports for the purpose of publication in accordance with EMA Policy 0070
- External guidance on the anonymization of clinical reports for the purpose of publication in accordance with EMA Policy 0070
- External guidance on the identification and redaction of commercially confidential information in clinical reports submitted to EMA for the purpose of publication in accordance with EMA Policy 0070. (EMA, 2018; EMA, 2019).

2.4.2 Anonymization and redaction

Redaction and anonymization of clinical data takes place before the documents are placed on the public website. Clinical data refers to all clinical documents and IPD (Individual Patient Data). Anonymized information refers to visible re-constructed data resulting in a low risk of individual identification, whereas redaction is a way of masking the data completely. As a rule, while redaction is the only option available for CCI, the authority prefers anonymization of PPD where possible. It should also be mentioned that Health Canada accepts submissions of redacted and anonymized packages that have previously been published under EMA Policy 0070. (EMA, 2018; Health Canada, Public Release of clinical Information: guidance document, 2019a).

2.4.3 Submission flowchart

In general, all submitted clinical documents of module 2.5, 2.7 and module 5 CSRs are to be published. In addition, other documentation must be included, such as justification tables and an anonymization report, and the dossier containing all these documents is called a submission package. (EMA, 2018).

EMA intends to publish the submission package 60 days after the European Commission decision of the MAA (EMA, 2021b). The sponsor sends the redaction proposal of the submission package through an e-submission gateway to EMA. The submission package contains the module 2 and CSR documents with proposed redactions, a table of justification for each document with a declaration text for CCI redactions and an anonymization report for PPD protection/redaction. The applicant will receive a receipt of the submitted package. Consultation takes place during the following 47 days. EMA reviews the justification tables, redaction proposals, anonymization report and sends comments to the applicant through the secure file transfer system, Eudralink. The applicant will then respond to the justification table comments made by EMA. In the final stage of conclusion process, a redaction conclusion notification is sent by EMA and a

consultation agreement is accepted by the applicant. During the following 27 days, after the redaction conclusion notification, the final redacted submission package is updated by the sponsor, including an updated cover letter and anonymization report. Simultaneously EMA ensures that the EPAR (European Public Assessment Report) is updated and the European Commission decision received. EMA will then send an acknowledgement receipt and submission acceptance email. The package is final when the clinical documents in question are redacted and accepted by the authority. A final version of the redacted document is now ready to be published and a watermark and document protection is added before the documents are placed on the public website (EMA, 2018, p. 16). This process is illustrated in Figure 8.

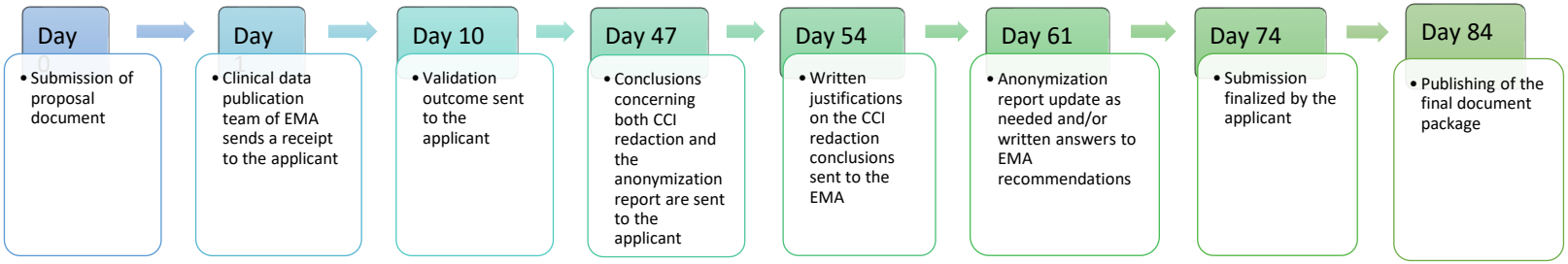

Figure 8. A process flowchart of EMA Policy 0070's end-to end process (EMA, 2018).

2.4.4 The consultation in detail

As soon as EMA has received the proposal document package, the task will be assigned to a team member. EMA performs a technical validation of the justification tables and a review of the anonymization report within 10 days. If applicable, the applicant must clarify the CCI approaches within 5-7 days. The consultation process is ensured via Eudralink. If clarification is requested but without an attached response, the proposal will be rejected. At the end, on day 47 at the latest, a final conclusion is sent to the applicant regarding the CCI evaluation and anonymization report recommendations or comments. If the anonymization report requires an update, a response is expected no later than day 61. EMA has 7 days to respond to the updated anonymization report. An agreement ought to be reached for the final redacted documents by day 74. See Figure 9. (EMA, 2018, pp. 31-32).

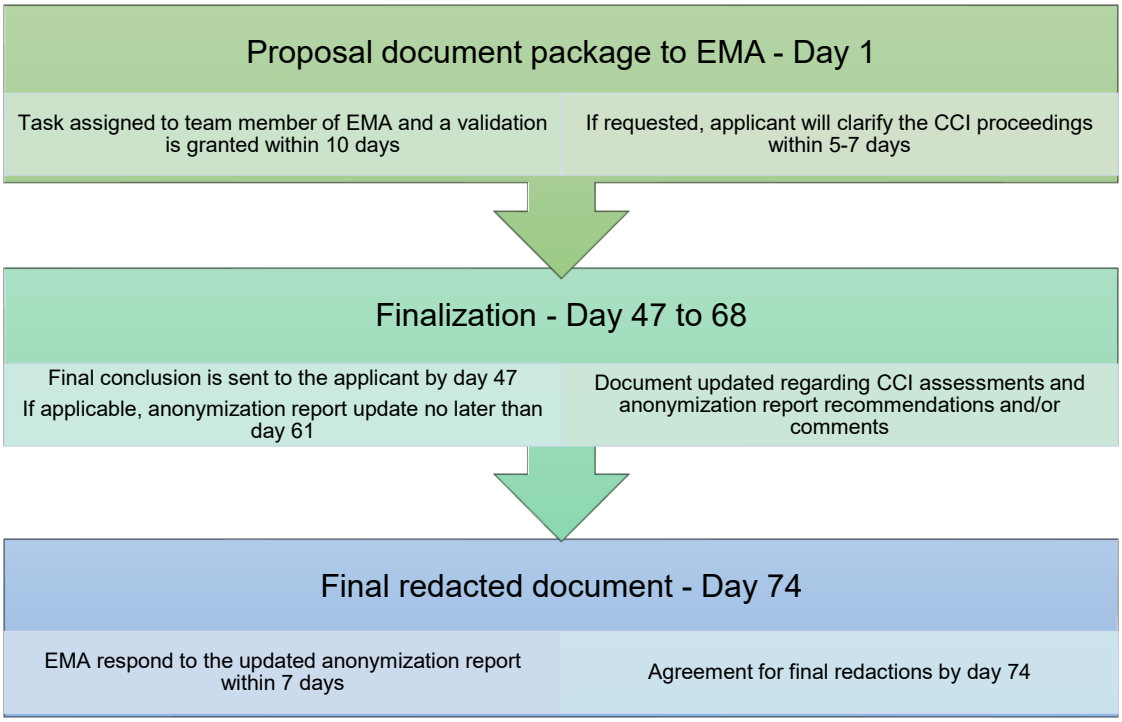

Figure 9. Consultation flowchart of redaction and anonymization (EMA, 2018).

2.4.5 Documents and application types

Documents in the final redacted package, submitted to EMA, are as follows (Table 5):

Table 5. Documents found in the final redacted package (EMA, 2018)

Cover letter	Justification tables	Clinical reports	Anonymization report
• Contains confirmation declarations saying that the submitted clinical reports are accurate • Not available for the public	• Used as a communication tool where the applicant clarifies the reasonings for CCI redactions in each document, while EMA evaluates the comments given • One justification table is submitted per document (if any CCI) • Not available for the public	• Module 2.5, 2.7 and CSRs • Submitted with the right redaction and anonymization approaches • For CCI the redaction is indicated by a black rectangular box marked with red text "CCI" • PPD redactions are indicated by a blue rectangular box marked with black text "PPD" • Published for the general public	• All the chosen PPD anonymization approaches are clearly explained by the sponsor • If details are unclear and further justification is needed, the applicant has the responsibility to reply with an updated version of the report • Published for the general public

CCI = Commercially Confidential Information; CSR = Clinical Study Report; PPD = Protected Personal Data

Different classification of drug applications is utilized depending on the circumstances. These are the application types in scope of Policy 0070 (Table 6):

Table 6. Classification of drug applications in scope of Policy 0070

Initial MAA	Article 58 application[7]	Extension of indication application	Type II variation
• Defined as a request for a new active substance **(European commission, 2015)**	• In collaboration with WHO, a scientific opinion is expressed and an article 58 describes the proposed reason for medicinal use outside of Europe (The European Union, 2004)	• When an application is applied to extend the indication used for a specific drug based on new clinical data • The extension applications vary depending on the applied changes • Extension applications are certain modifications in the drug such as changes in the active substance, in strength, route of administration or drug form (The European Union, 2008)	• If a major variation is implemented that might affect the result of quality, safety or efficacy (The European Union, 2008)

WHO = the World Health Organization

[7] Article 58 application is to facilitate patient access to essential drugs in low- and middle-income countries for diseases of major public health interest (EMA, n.d. (b))

2.5 Health Canada PRCI

Health Canada released, in the beginning of 2019, a corresponding guidance to EMA Policy 0070, called Health Canada PRCI. These regulations use Canada-specific terminology and the PRCI abbreviation stands for Public Release of Clinical Information. Health Canada PRCI is the reason for the opportunity to download module 2 documents from the clinical information portal of Health Canada. (Health Canada, 2019a).

2.5.1 Brief introduction

Health Canada PRCI guidance is available to provide instructions on how to manage anonymization and redaction processes when applying for marketing authorization of drug packages in Canada and after this publishing the documents online. This guidance has been created with the awareness of the Privacy Act of Canada - a federal legislation from 1983 amplifying how personal information is to be dealt with. (Health Canada, 2019a).

Compared to EMA Policy 0070, PRCI covers both drug submissions and medical device applications. In addition, some differences in wording exist, e.g., EMA addresses confidential information as commercially confidential information (CCI), while HC's corresponding terminology is confidential business information (CBI) (Health Canada, 2019a). Likewise, protected personal data (PPD) in the EMA regulation is comparably defined as personal information (PI) in HC guidelines (Office of the Privacy Commissioner of Canada, 2019).

Equivalent anonymization and redaction rules, mentioned in section 2.4.2, apply for Canada (Health Canada, 2019a).

2.5.2 Submission flowchart

The public release of clinical information is divided into five phases and should be completed within 120 days of the process start (Figure 10). When a submission review of a drug package has started and is accepted, the applicant will receive a notice of the PRCI process. (Health Canada, 2019a).

This notice will request the applicant to redact the documents that will be published. These documents should be modified within 60 days to comply with the editing rules set by Health Canada PRCI. During this time, the applicant also has an opportunity to initiate a meeting (process initiation meeting, PIM) with HC where the applicant can discuss and clarify requirements regarding PRCI. This meeting should take place between 120 days before HC makes its final decision on the submitted drug package and 20 days after this decision is made. A redaction CBI control sheet, equivalent to EMA's justification tables, should be included in the submission package together with the PI anonymization report. The proposed redactions will be reviewed by HC and the applicant's proposals will be either accepted, partially accepted or rejected. An alternative option is to request a submission of previously redacted documents approved by EMA. The final package should be submitted for publication via the Common Electronic Submission Gateway (CESG) and thereafter, according to the transparency guidelines, it will be openly published on HC's clinical information website, for all to read. (Health Canada, 2019a).

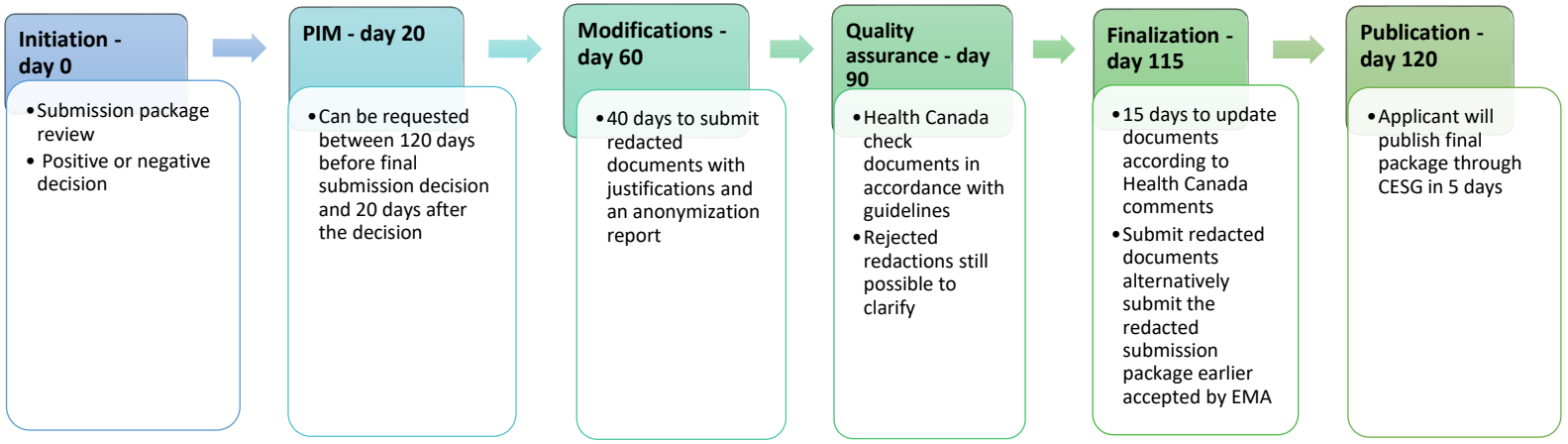

Figure 10. A process flowchart of Health Canada PRCI's end-to end process (Health Canada, 2019a).

2.5.3 Documents and application type

Documents in the final redacted package, submitted to Health Canada, are as follows (Table 7):

Table 7. Documents found in the final redacted package (Health Canada, 2019a)

Redaction control sheet	Clinical reports	Anonymization report
• Template excel sheet as communication tool • The applicant clarifies the reasonings for CBI redactions in each document, while HC evaluates the comments given • Not available for the public	• Module 2.5, 2.7 and CSRs • Submitted with the right redaction and anonymization approaches • Redaction is achieved by covering the text with a rectangular box • CBI and PI information must be clearly disguisable • Published for the general public	• All the chosen PI anonymization approaches are clearly explained by the sponsor • If details are unclear and further justification is needed, the applicant has the responsibility to reply with an updated version of the report • Published for the general public

CSV = Comma-Separated Values; CBI = Confidential Business Information; HC = Health Canada; CSR = Clinical Study Report; PI = Personal Information

Similar to EMA, HC has several application categories that are in scope of the PRCI, as applicable. The wording, however, is quite different from the EMA guidelines and is as follows (Table 8):

Table 8. Health Canada application types (Health Canada, 2019a)

NDS	NDS-NAS	SNDS-c	Rx-switch
• New Drug Submission • Active substances that have previously been classified • For NDS application all necessary information must be provided from start to finish, a full process review	• New Drug Submission - New Active Substance • No earlier variations of other previously approved active ingredients in Canada	• Supplemental New Drug Submission holding confirmatory trials	• The submission to ask a drug status to change from prescription drug to over-the-counter drug
SNDS	**SANDS**	**ANDS**	**SEUNDS**
• Supplement to a New Drug Submission • Applications for changes in packaging, labeling, dosage, ingredients or new indication, for a previously submitted NDS • Must show data of safety and efficacy of implemented changes	• Supplement to Abbreviated New Drug Submission • Applications for changes in packaging, labeling, dosage, ingredients or new indication, for a previously submitted ANDS • Must show data of safety and efficacy of implemented changes	• Abbreviated New Drug Submission • The submission to obtain marketing approval for a generic product • Safety and efficacy data equal to the original drug	• Supplemental Extraordinary Use New Drug Submissions

HC has divided their transparency implementations into four stages: year 1, year 2, year 3, year 4. In year 1, NDS-NAS, SNDS-c, Rx-switch submissions and COVID-related (IO/SNDS) are published. In year 2, proactive publication of clinical information is valid for all NDS (meaning both new active substances and those which do not fall under this category), SNDS-c and Rx-switch. For year 3, all SNDS are to be proactively published. For instance, these can be applications for new indications of an existing product. Year 4 involves implementing ANDS, which refers to generic drugs. HC also implements medical devices in year 3 and 4, which is not brought up as essential information in this thesis. Note that some steps are yet to be implemented and the different stages, referred to as year 1 to 4, are not equal to calendar years. All drug submissions (and device applications) which already have final regulatory decisions remain subject to release on request. (Health Canada, 2019a; Bayer, Personal Communication, 2021).

3 Aims

3.1 Goals

The thesis intends to measure sponsors' module 2.5 and 2.7 documentation and writing approaches from a medical writing standpoint and, hopefully, find guiding tools to proceed with a simplification process, while ensuring that relevant content is still provided. However, it is important to remember that the writing strategies depend on both the sponsor's approaches and the document's structure, which may vary due to different indications, active substances and excipients. As certain details are impractical to compare between sponsors, this thesis is designed to investigate on a more surface level and to create a base for future research.

The purpose is to clarify whether extensive variations between sponsors' writing approaches appear and whether something can be learned from these analyses. The current standards for document writing and information redaction are remarkably time-consuming and require numerous resources. It would be beneficial to be able to eliminate redundant and repeated information in the documents.

The long-term goal is to reduce time, improve the writing quality and minimize the redaction requirement. The aim is to make every sponsor that is dealing with MAAs aware of the simplification suggestions developed from the findings and of how they can be implemented in their daily work.

3.2 Objectives

The two main objectives for identifying writing simplification possibilities in publicly available module 2 documents are presented in Figure 11.

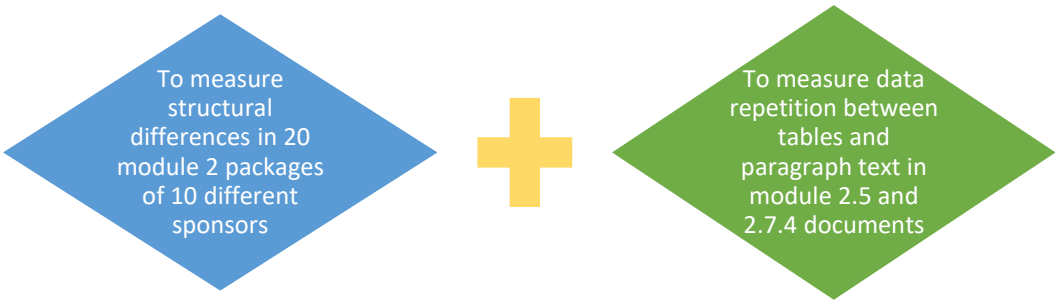

Figure 11. The main goals and research questions for the research project.

The objectives in Figure 11 are achieved by evaluating publicly available module 2 documents from submission packages published as part of EMA Policy 0070 and HC PRCI processes, and prepared by ten different pharma sponsors. The following questions are answered in the results, conclusion and discussion to support the objectives:

- What were the structural differences in the 20 module 2 packages from the different sponsors?
- How well is the CTD guideline followed?
- What are the possible simplification suggestions?

The author aims to clarify the objectives by analyzing what kind of variations are found in the collected results. As a conclusion, the results are further compared to the CTD guideline to investigate whether all sponsors are following the recommendations. Moreover, possible simplification suggestions will be prepared for writing of module 2 documents based on the findings.

4 Methodology

The research is divided into two parts: an independent part and a collaborative part.

The independent work concerned downloading the chosen documents. The documents were viewed, quantitative data were collected, and finally a comparison of the writing strategies and document structures were processed. The author determined that the most convenient way to collect and compare data to conduct the research was by following the approach described in section 4.3.

The collaborative part involved a global expertise team who discussed, analyzed and evaluated the collected variable data. The collaboration team also discussed additional ways to contribute to an industry-broad recommendation for writing module 2 documents.

The validity and reliability of this research will undergo substantial comparison, evaluation and discussion with senior team members. As this is a pioneer research subject, the trustworthiness will depend on the outcome of this research followed by future research findings. The research activities are described in Figure 12.

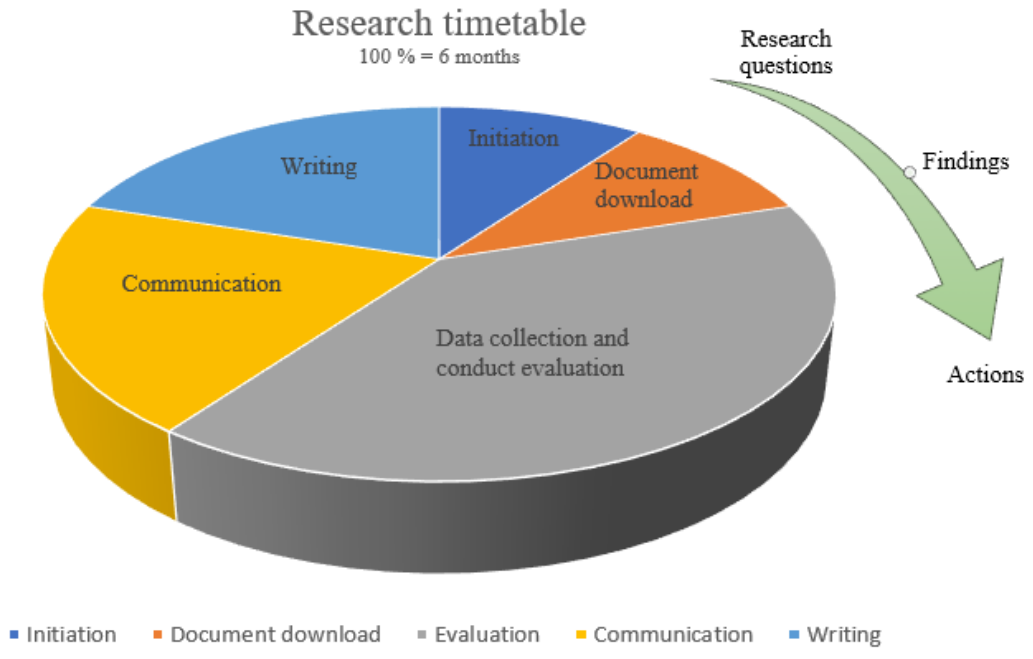

Figure 12. A 6-month research schedule divided into different topics.

A timetable has been created to delineate the total workload. Initiation meetings were held to hone the research questions, and document downloading could then begin after careful preparation. After downloading the desired documents, the evaluation was

conducted. The author's findings were then communicated with the rest of the team and actions towards a documentation simplification process were initiated.

4.1 Literature review

In order to have a better understanding of the module 2 documents, a literature review of the CTD, MAA process, and EMA and HC transparency guidelines was conducted. The CTD describes the rules concerning document structure that are essential for the MAA, and the MAA process clarifies the road from an investigational drug to market access and the transparency guidelines clearly state what is meant by transparency, its necessity and how to apply this within the documents.

4.2 Documents and sponsors

One submission package may include hundreds of documents and thousands of pages, which is too much data for a master's thesis evaluation. Therefore, the scope of this thesis research solely concerns publicly available module 2 documents.

To be able to download the module 2 documents, the author registered in the EMA web portal. Downloading documents from the HC portal only requires the reader to agree with the terms of use, while EMA requires passport or ID information to be entered as well as a reason to be stated for full access. There are two types of registration for accessing the EMA portal: standard and academic. Only an academic registration allows full access, enabling documents to be downloaded or searched for specific words. After registering, the author was allowed to download the published module 2 documents.

In this thesis, one module 2 package refers to the publicly available documents of the five modules: 2.5, 2.7.1, 2.7.2, 2.7.3, 2.7.4. Altogether 20 module 2 packages were downloaded from EMA and HC web portals and the packages are from ten different pharma sponsors. The sponsors were randomly selected, however, each sponsor chosen had to have a published module 2 package in both portals. The reason for evaluating more than one package per sponsor is to understand documentation differences within a single sponsor.

Because some EMA activities have been suspended and new submission packages are on hold (excluding Covid-19 products), the author has concluded that the most valid

option is to download one module 2 package from the EMA portal and the other module 2 package from the HC portal, for every sponsor. The HC portal contains the latest submission packages and is, therefore, important when investigating the document structures of each sponsor for the document simplification initiation, while EMA is limited to somewhat older submission packages with possibly outdated writing strategies.

The scope of document-type examination is limited to initial MAA packages for EMA, and NDS-NAS and NDS packages for HC (see sections 2.4.5 and 2.5.3). The lack of variability when choosing the submission packages from corresponding portals must be highlighted, especially when emphasizing new drug substances. This is due to the newly launched portals and the fact that document downloading of submission packages has just come to life.

4.3 Variables

At the time of the document downloading, the following information was recorded:

- Indication, first-in-class[8], publication year and document type

The following variables were chosen for quantitative evaluation and comparison:

- Number of pages
- Number of tables[7]
- Number of figures
- Number and structure of level-1 headings and subheadings
- Categorization of table[9] content repetition in the text

The variables are measured as differences within the documents and means and averages are calculated to create tables and figures to visualize how writing strategies differ between sponsors and different modules.

The number of pages is calculated from the first page of the document to the last page of the text body, this excludes the reference list and any appendices from the calculation. The number of tables per document is considered specifically according

[8] A drug that uses an innovative mechanism of action for treating a disease

[9] A table is referred to as a table within the text paragraphs of the module 2 document body. Tables do not usually present the full data from a statistical output table which is located either after the text sections (as tables, listings and figures), in section 14 in CSRs or in a separate integrated analyses document

to the table of tables, in the same way that the number of figures is calculated according to the table of figures.

The level-1 headings are the first-grade headings provided by the CTD ICH guidelines (see Table 2), excluding conclusion, reference list and appendix. The analysis intends to show whether sponsors follow the CTD structure for level-1 headings as well as how reference lists and appendices are provided for each sponsor.

Headings grouped under the level-1 headings are called subheadings. The number of CTD-specific subheadings has been calculated to estimate how much text these subheadings contain. In addition, the number of pages that belong to the level-1 headings as per the CTD structure has been calculated to form a subheading-page count ratio. This allows the analysis of how many subheadings are provided per module for each document, how much text is provided per subheading and how detailed the documents are.

Additionally, content repetition between tables and the text body was measured by manually reviewing and categorizing the review results into self-set categories, which are presented in more detail in Table 9. As this activity was shown to be quite time-consuming, the analysis was only done for module 2.5 and 2.7.4 documents.

Table 9. Explanations and examples of the categorizing system of table repetition in the text

Category	Detailed explanation
1. Data from table is not repeated in text	Nothing or only the title is mentioned from table in the text. Example: • See table x • The title a is presented in table x
2. Data from table is summarized in text	A general overview of the table, where no data from the table is repeated. Example: • The most common AEs are found in table x • The most common AEs are a, b and c • All the study arms of study r are shown in table x
3. Data from table is somewhat repeated in text	An overview or thorough presentation of the table data in the text where specific data are repeated Example: • Table x shows that y% of the subjects experienced one or many AEs • X number of the subjects were withdrawn from the study
4. Data from table is (almost) completely repeated in text	All data or almost all data are repeated from the table in the text, leading to full repetition of data that has already been presented Example: • Table x shows that u% experienced AE a; p% AE b, s% AE c …

AE = Adverse Event

The categories depict how various sponsors choose to present their data and if repetition is common in the documents. Notice that not all tables present data in numerical form, but some tables present data as text. Both numerical and text data are subject to the same categorizing system.

4.4 Discussion group

A discussion group of five experts was established to support the author throughout the process. Deliberately a mixed group of senior experts, working in different functions with different scientific backgrounds, was set up to allow wider competency and lively discussions about the author's findings. A monthly meeting was initiated to ensure everyone stayed up-to-date throughout the whole process. Additional meetings were arranged, and more experts were consulted as needed.

4.5 Coding of sponsors and packages

The goal of the research is to compare the set variables in an anonymous manner. Therefore, the author has decided to create a coding system unique for this research that allows the reader to differentiate between companies and their submitted drug packages without revealing company information (Table 10). After all, this research intends to focus on the industry rather than any specific sponsor.

Table 10. Coding strategy to keep sponsors and their drug packages anonymous

				European Medicines Agency	*Health Canada*
Company name	Sponsor	Package code 1	Package code 2	Submission package 1	Submission package 2
Anonymous	A	A-1	A-2	Anonymous	Anonymous
Anonymous	B	B-1	B-2	Anonymous	Anonymous
Anonymous	C	C-1	C-2	Anonymous	Anonymous
Anonymous	D	D-1	D-2	Anonymous	Anonymous
Anonymous	E	E-1	E-2	Anonymous	Anonymous
Anonymous	F	F-1	F-2	Anonymous	Anonymous
Anonymous	G	G-1	G-2	Anonymous	Anonymous
Anonymous	H	H-1	H-2	Anonymous	Anonymous
Anonymous	I	I-1	I-2	Anonymous	Anonymous
Anonymous	J	J-1	J-2	Anonymous*	Anonymous

**HC submission*

The table above explains how each sponsor or company received an anonymous code from A to J. The author has then created a relationship between each sponsor and its chosen module 2 documents. For instance, sponsor A has two submission packages that are to be evaluated, one from EMA (A-1) and one from HC (A-2). Each number stands for one specific document. Note that due to lack of documents within the research's scope, both submission packages were exceptionally downloaded from HC for sponsor J.

5 Results

The objectives of the thesis are described in section 3.2 and shortly below.

- To measure structural differences in 20 module 2 packages from 10 different sponsors
- To measure data repetition between tables and paragraph text in module 2.5 and 2.7.4 documents

The above objectives will aid in evaluating suggestions for possible documentation simplification.

5.1 Basic characteristics of the MAAs

Table 11. presents characteristics of the documents that can be found in the respective portals for each MAA. These details were collected during the document download and display the following information: indication, whether it is a first-in-class drug (meaning a drug that uses an innovative mechanism of action for treating a disease) or not, the year of document publication and the document type. The collected information forms the baseline and allows to understand the similarities and differences between the document packages at the start.

More than half of the documents (60%, 12/20) were oncology indications, documents A-2 and D-1 presented RA (Rheumatoid Arthritis) indication and six of the documents (B-2, C-2, E-1, F-1, H-1 and H-2) were for a variety of indications. Of all the drugs, 35% (7/20) were categorized as first-in-class drugs. No MAA was published before the year 2016 and most applications were published between 2018 and 2020. All MAAs for EMA were subject to the "initial marketing authorization" procedure and for Health Canada, "NDS" or "NDS-NAS" procedures.

Table 11. Specific characteristics of each document. Each color represents a specific sponsor and its specific document

Sponsor	*Indication*	*First in class*	*Publication year*	*Application type*
A-1	Oncology	Yes	2018	Initial marketing authorization
A-2	Rheumatoid Arthritis	No	2020	NDS-NAS
B-1	Oncology	No	2016	Initial marketing authorization
B-2	Osteoporosis	Yes	2019	NDS
C-1	Oncology	Yes	2017	Initial marketing authorization
C-2	Hyperkalemia	No	2020	NDS
D-1	Rheumatoid Arthritis	No	2018	Initial marketing authorization
D-2	Oncology	No	2019	NDS
E-1	Antiviral	No	2018	Initial marketing authorization
E-2	Oncology	No	2020	NDS
F-1	Infection profylax	No	2017	Initial marketing authorization
F-2	Oncology	No	2020	NDS
G-1	Oncology	Yes	2017	Initial marketing authorization
G-2	Oncology	Yes	2020	NDS-NAS
H-1	Diabetes mellitus	No	2018	Initial marketing authorization
H-2	Haemophilia	No	2019	NDS
I-1	Oncology	Yes	2018	Initial marketing authorization
I-2	Oncology	No	2020	NDS
J-1	Oncology	No	2020	NDS
J-2	Oncology	Yes	2020	NDS

The three following figures are presented below: number of pages (Figure 13), total number of tables (Figure 14) and total number of figures (Figure 15). The x-axis presents all publicly available modules and the y-axis presents the dependent variable (number of pages, tables or figures). Each document is displayed with a specific color.

5.2 Number of pages

The number of pages for each module (Figure 13 and Table 12) varied between sponsors.

The shortest module 2.5 document (F-1) consisted of 26 pages, while the longest document (A-2) was 145 pages. In other words, there was a 119 pages difference between the module 2.5 document extremes. However, the majority of the documents (95%, 19/20) had at least 47 pages and the mean was 79 pages and the median 80 pages.

There were fewer variations for module 2.7.1 documents and the number of pages was 59 or lower for all, except for one document (I-1) that had 105 pages. The shortest document (E-2) had 3 pages, followed by a document (F-1) with 5 pages. The difference between the module 2.7.1 document extremes was 102 pages. The majority of the documents (70%, 14/20) had at least 20 pages and the mean number of pages was 35 and the median 29 pages.

In module 2.7.2 documents, the longest document (J-1) had 195 pages, whilst the shortest document (F-1) had 12 pages. There was an increased difference between the module 2.7.2 document extremes, 183 pages, but 50% (10/20) of the documents had 100 pages or more. The mean number of pages was 101 and the median 100.

In module 2.7.3 documents, the longest document (A-2) had 210 pages and the shortest document (F-1) contained 15 pages. The difference between the module 2.7.3 document extremes increased further, reaching 195 pages. However, 45% (9/20) of the documents had at least 100 pages. The mean number of pages was 100 and the median 93.

The documents with peak number of pages were found in module 2.7.4 documents where two documents exceeded 400 pages, being 406 (A-2) and 407 (D-1) pages long. The shortest document (F-1) contained 12 pages. The largest difference between the extremes, 395 pages, was found in module 2.7.4 documents. Of all the 2.7.4 documents, 40% (8/20) amounted to at least 200 pages. The mean was 178 pages and the median 185 pages.

The longest document for the total clinical summary document, meaning module 2.7.1, 2.7.2, 2.7.3, 2.7.4 combined, contained 792 pages, while the shortest document was 44

pages long. The difference between the extremes was 748 pages. Of the total clinical summary documents, 45% (9/20) had more than 400 pages.

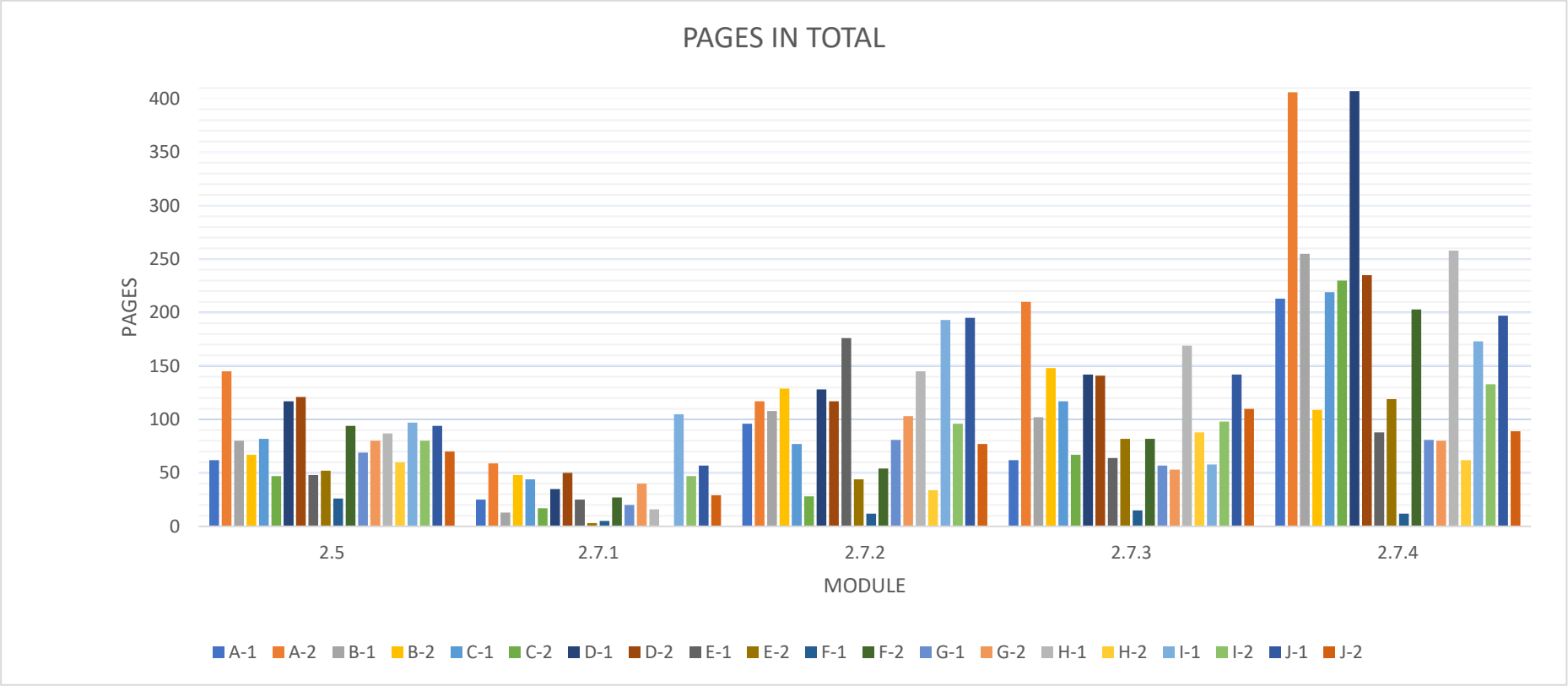

Figure 13. Total number of pages per document. Each color represents a specific sponsor and its specific document.

Table 12 shows that module 2.7.1 documents contained the lowest number of pages for all descriptive statistics (mean, median, max and min). The trend was followed by module 2.5 documents, except for the minimum number of pages, which was the highest of all the modules. The statistical values were similar between module 2.7.2 and 2.7.3, and module 2.7.4 documents had the highest statistical values for all, except for the minimum number of pages.

Table 12. Descriptive statistics of the pages in total (mean, median, max and min) for each document module plus a statistics combination of module 2.5 and 2.7 documents

Module	2.5	2.7.1	2.7.2	2.7.3	2.7.4	2.7 - total clinical summary document	2.5 + 2.7 combined
Mean	79	35	101	100	178	104	99
Median	80	29	100	93	185	88	82
Max	145	105	195	210	407	407	407
Min	26	3	12	15	12	3	3

5.2.1 What were the structural differences in the 20 module 2 packages from the different sponsors?

A relationship between the length of a document and the different modules could be observed. For example, document A-2 (RA indication) and J-1 (oncology indication) had a greater number of pages and stayed above average in every module, whilst document E-2 (oncology indication) maintained a lower page count for all modules.

Some exceptions were observed in the collected data. Document F-1, written for infection prophylaxis, and H-2, written for haemophilia, had exceptionally short documents in every module. Document F-1 did not exceed 30 pages for any module and H-2 also stayed well below average for all modules and had no document for module 2.7.1. However, sponsor H and F did not show a similar trend for documents H-1 and F-2, hence, a sponsor-specific approach to have shorter documents was not spotted and the indication might have affected the length of the document. Document A-2 and D-1 had especially long documents in module 2.7.4, the only documents for RA indication. Though, as this was an initial analysis, there is no possibility to draw a conclusion between indications and document structure. As the number of analyzed documents are especially short, it is believed that this finding was based on a coincidence.

It was apparent from a few documents that the sponsors had applied a marketing authorization for their product with a lower than average number of pages, however, this was not sponsor-specific as one sponsor could have submitted one short and one long document.

5.2.2 How well is the CTD guideline followed?

It is mentioned in the CTD guideline that the module 2.5 clinical overview documents should include around 30 pages, but this might vary depending on the complexity of the application. The data collected for the clinical overview documents showed that only one document (F-1) abided by the recommended number of pages, whereas the rest of the documents (95%, 19/20) exceeded the recommendation. In other words, most sponsors did not follow the recommended document length of 30 pages.

According to the CTD guideline, the target length of the total clinical summary document 2.7 (modules 2.7.1 to 2.7.4 combined) ranges from 50 to 400 pages. The CTD guideline for the recommended length was not followed as the length of the module 2.7 documents was exceeded by 45% (9/20) of the documents.

The CTD guideline encourages to minimize the length by including graphs and tables in the text body. Additionally, it is not recommended to repeat material in the module 2.5 clinical overview that has been mentioned elsewhere, hence, cross-referencing is a preferred solution (EMA, 2004). Cross-references into module 2.7 and module 5 would be the most beneficial approach for module 2.5 to avoid unnecessary repetition and reduce the document length. Cross-referencing into module 5, whenever applicable, could decrease the length of module 2.7 documents. Furthermore, electronic CTD admits the generation of active cross-references, and this allows easy access to the referenced data with one click.

5.3 Number of tables

The total number of tables (Figure 14 and Table 13) showed a growing trend from module 2.5 to module 2.7.4 documents.

For module 2.5 documents, the lowest number of tables was 4 tables (F-1), while the document with most tables had 33 (G-2). The difference between sponsor extremes was 29 tables. The mean was 14 tables and the median 12 tables.

For module 2.7.1 documents the highest number of tables was 46 (I-1) and the lowest was 0 (E-2), and the difference between the sponsor extremes was 46 tables. The mean and median number of tables was 15 for both. Document H-2 did not include any document for module 2.7.1.

For module 2.7.2 documents, the maximum number of tables was 68 (J-1) and the minimum 2 (F-1). The difference between sponsor extremes was 66 tables. The mean was 29 tables and the median 26 tables.

The highest number of tables for module 2.7.3 documents was 71 (J-2) and the lowest was 2 (F-1). The difference between sponsor extremes was 69 tables. However, out of all the modules, module 2.7.3 documents fluctuated the least between the sponsors and, for most documents (60%, 12/20), the number of tables stayed between 10 and 30. Module 2.7.3 documents had a mean of 30 tables and the median was 26 tables.

The most extensive variations were seen in module 2.7.4 documents where the largest number of tables was 145 (A-2) and the lowest was 1 table (F-1). The difference between the extremes was 144 tables. However, 85% (17/20) of the documents had 80 or fewer tables. The mean was 59 tables and the median 60 tables.

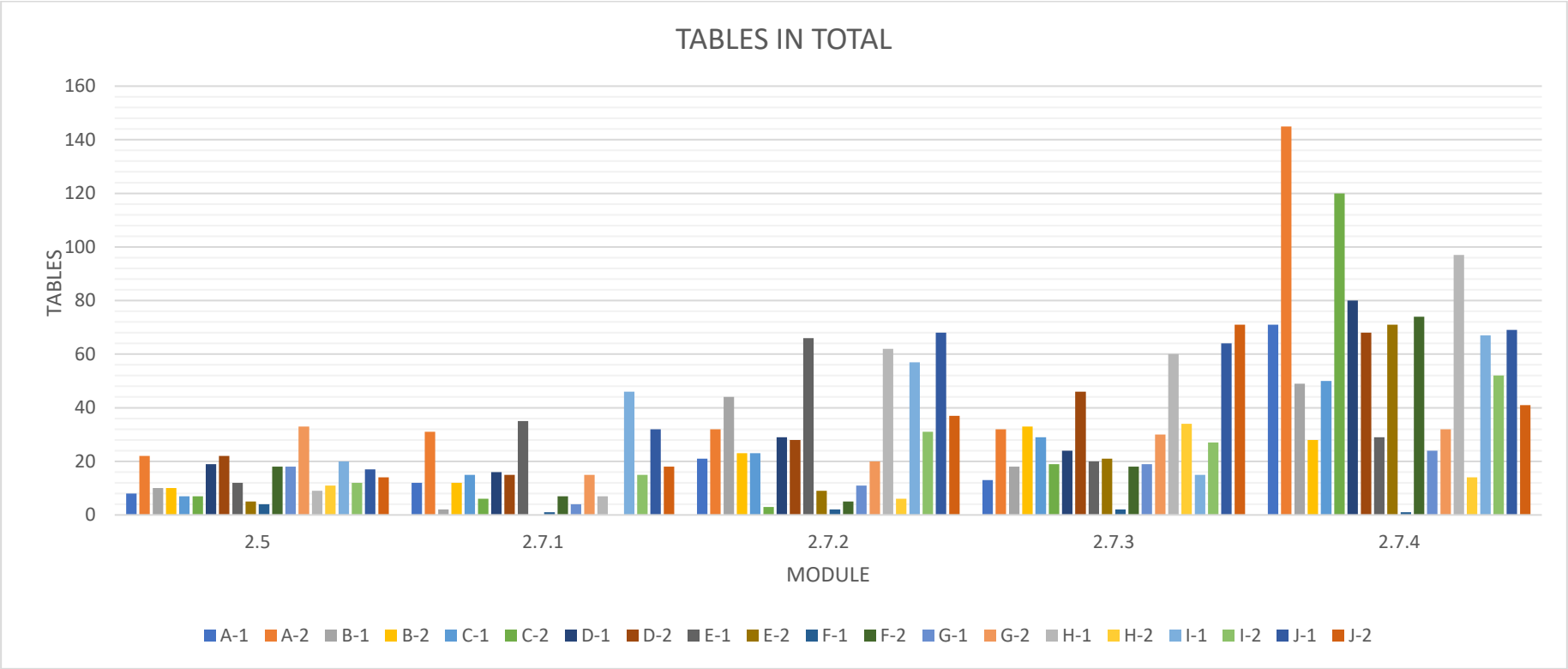

Figure 14. Tables in total for all document modules. Each color represents a specific sponsor and its specific document.

Table 13 shows that module 2.5 documents contained the lowest number of tables for all descriptive statistics (mean, median, max and min) followed by module 2.7.1 documents. The statistical values were similar between module 2.7.2 and 2.7.3, whilst module 2.7.4 documents had the highest statistical values for all, except for the minimum number of tables.

Table 13. Descriptive statistics of the tables in total (mean, median, max and min) for each document module plus a statistics combination of module 2.5 and 2.7 documents

Module	2.5	2.7.1	2.7.2	2.7.3	2.7.4	2.7 (the total clinical summary document)	2.5 + 2.7 combined
Mean	14	15	29	30	59	33	30
Median	12	15	26	26	60	28	21
Max	33	46	68	71	145	145	145
Min	4	0	2	2	1	0	0

5.3.1 What were the structural differences in the 20 module 2 packages from the different sponsors?

For the majority of sponsors, the number of tables was growing, from module 2.5 towards module 2.7.4 documents (Figure 14). Three documents (A-2, H-1 and I-1) had a greater number of tables in at least one of the modules, which clearly correlated with the total page count of the specific document. For example, A-2 had 145 tables (highest digit for the module) and 406 pages (highest digit was 407) in the module 2.7.4 document, H-1 had 62 tables and 145 pages in the module 2.7.3 document and I-1 had 46 tables and 105 pages in the module 2.7.1 document. However, this was not true for all documents, for example, C-2 had 120 tables for the module 2.7.4 document (the second highest digit of all module 2.7.4 documents) but only 230 pages (the maximum number of pages for module 2.7.4 was 407 pages).

Some exceptions were observed in the collected data. Documents of F-1 had an exceptionally low number of tables in every module and document A-2 had an exceptionally high number of tables in module 2.7.4. Document of E-2 did not present any tables in module 2.7.1.

5.3.2 How well is the CTD guideline followed?

Per the CTD guideline, tables are always preferred in the content if it can improve the readability of the document, but it is up to the sponsor to decide when information and

results should be presented as tables, figures or text. This is a quite vague statement, which has resulted in different interpretations, as shown by the collected data in Figure 14 and Figure 15.

The lowest number of tables stayed similar between every module (between 0 and 4), whilst the maximum number of tables increased consecutively from module 2.7.1 to 2.7.4 documents (from 46 to 145 tables). The module 2.5 documents, which are supposed to be short clinical overview documents, had the lowest number of tables (maximum 33 tables). The requested content for module 2.7.1 documents is also fewer, hence, the number of descriptive tables should be few. On the other hand, the recommended content for module 2.7.4 documents is extensive and should contain more detailed descriptions., thus, more tables are to be provided in this document. This was true for all modules of the collected data. However, as the documents exceeded the recommended page length for both modules (2.5 and 2.7), it could be discussed whether all provided tables are essential.

5.4 Number of figures

For the total number of figures (Figure 15 and Table 14) no clear trend between the modules was shown. In module 2.5, two documents (E-1 and F-1) did not include any figures and the largest number of figures was 21 for document D-2. The difference between the sponsor extremes was 21 figures. The mean for the number of figures was 9 and the median was 8.

In module 2.7.1 no figures were presented in six of the documents (C-2, E-1, E-2, F-1, G-1 and H-1). H-2 did not contain any document for module 2.7.1, hence, no figures were presented for this document. The highest number of figures was 16 and was recorded in document I-1. The difference between the sponsor extremes was 16 figures. The mean and median were 6 figures for both.

The documents in module 2.7.2 contained the most figures, up to 52 (B-2 and H-1), yet one document (F-1) did not include any figures. Module 2.7.2 documents also showed the largest fluctuation between sponsors and the difference between them was 52 figures. The mean was 25 figures and the median was 21 figures.

Module 2.7.3 and 2.7.4 had a maximum of 38 and 44 figures in their documents, with a considerably lower mean of 16 and 10 figures, respectively. In addition, one document (F-1) for module 2.7.3 and four documents (C-2, E-2, F-1 and I-1) for module 2.7.4 contained no figures.

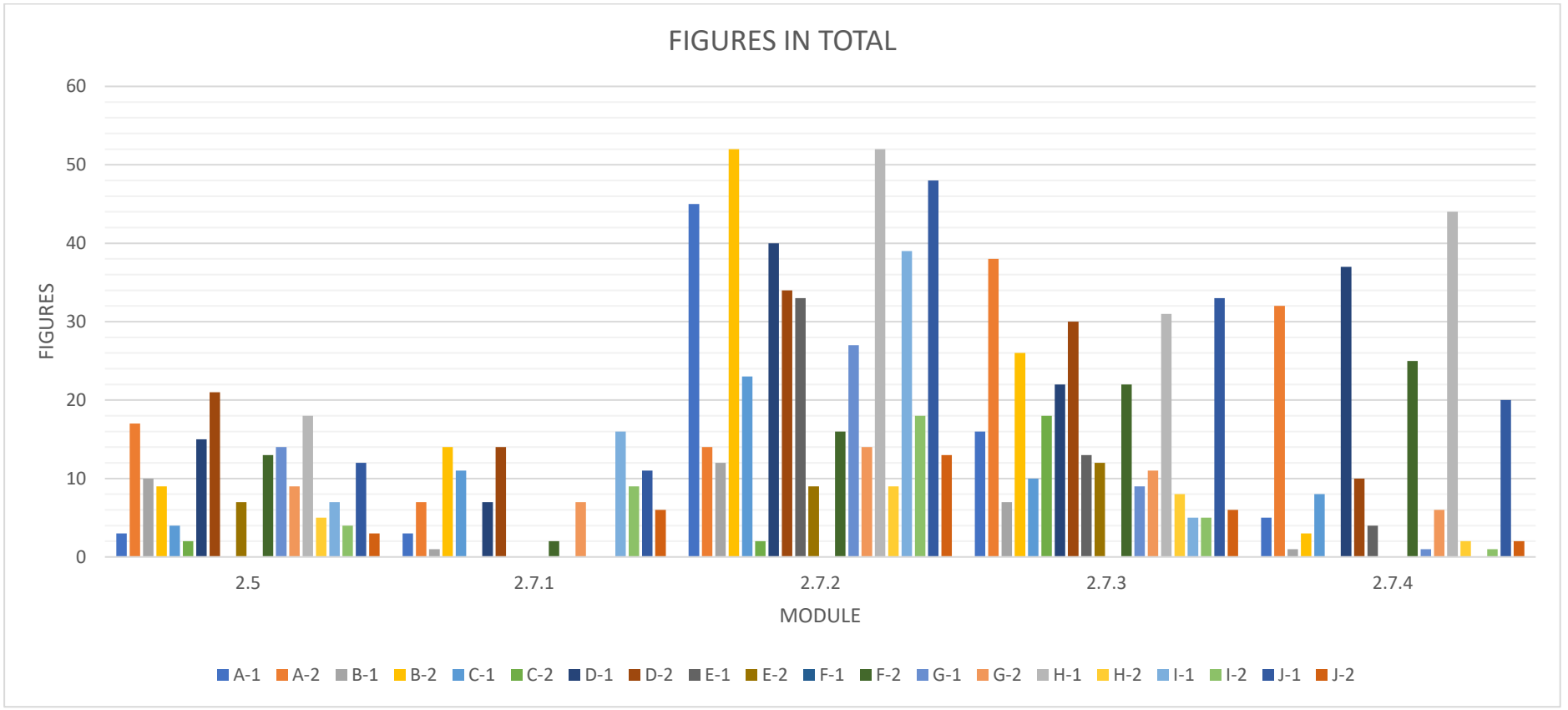

Figure 15. Total number of figures for each document module. Each color represents a specific sponsor and its specific document

Table 14 shows that module 2.7.2 documents had the highest values for the number of figures for the different descriptive statistics (mean, median, max and min). The minimum number of figures was 0 for all modules.

Table 14. Descriptive statistics of the figures in total (mean, median, max and min) for each document module plus a statistics combination of module 2.5 and 2.7 documents

Module	2.5	2.7.1	2.7.2	2.7.3	2.7.4	2.7 (the total clinical summary document)	2.5 + 2.7 combined
Mean	9	6	25	16	10	14	13
Median	8	6	21	13	4	10	9
Max	21	16	52	38	44	52	52
Min	0	0	0	0	0	0	0

5.4.1 What were the structural differences in the 20 module 2 packages from the different sponsors?

As presented in Figure 15, a large fluctuation in number of figures between both sponsors and the modules was noticed. The most infrequent use of figures was shown in the module 2.7.2 and 2.7.4 documents. The infrequency in module 2.7.2 is unclear but might depend on how the sponsor chooses to present PK data (in figures or tables). For module 2.7.4 documents, 75% (15/20) presented 10 or fewer figures. The remaining five documents (A-2, D-1, F-1, H-1 and J-1) presented a greater number of figures (between 20 and 44) and, in addition, these documents were long with number of pages ranging from 197 to 407. Another reason for the infrequent use of number of figures in module 2.7.4 might be the large difference between the total page count extremes (395 pages).

Some exceptions were also observed in the collected data. There were large variations between a single document for different modules. Some documents had reached the average number of figures for certain modules, while only a few or no figures at all were included in other modules. It was observed that some documents (e.g., A-2, H-1) had implemented more figures in all the modules, compared to other documents (H-2 and J-2). Document F-1 did not include figures for any module and, as shown earlier, the same document also had fewer pages and tables. The reason for this might be that even though it is an initial marketing authorization product, the active substance itself is known from before and, therefore, the information presented in the documents are reduced as enough data are already available.

5.4.2 How well is the CTD guideline followed?

As stated earlier, figures are always preferred in the content if it can improve the readability of the document, but it is up to the sponsor to decide when information and results should be presented as figures.

For the total number of figures in the documents, it is shown that there is no clear guidance on how often figures should be used. As mentioned above, PK data, which present measurements concerning the concentration and duration of the drug in the blood, is found in module 2.7.2 documents. These parameters can describe bioavailability, metabolism, duration, distribution and excretion, and the data are often better presented as figures rather than tables.

5.5 Level-1 headings

The following five tables (Table 15 to Table 19) were created to examine whether the sponsors followed the recommended CTD standards for all level-1 headings for the publicly available module 2 documents and if any structural difference appeared between sponsors. Another topic evaluated was how reference lists and appendices are provided in the documents by each sponsor. The detailed information of the module structures and level-1 heading naming can be found in section 2.1, Table 2.

The first column for each table shows the sponsors and their specific document, and the second column describes whether an executive summary was included in the beginning of the document. The executive summary is a short overview of what will follow in the text body. The following columns after the executive summary display the level-1 headings that vary in number depending on the given module (2.5, 2.7.1, 2.7.2, 2.7.3 or 2.7.4). Any additional level-1 headings (meaning level-1 headings that are not CTD recommended) are sponsor-specific[10] and mentioned in a column of their own. The conclusion, reference list and appendix are mentioned in the tables as separate level-1 headings if applicable.

Table 15 shows the structure of the clinical overview module 2.5 document. An executive summary was included in eight (40%) clinical overview documents. All but two (E-1 and E-2) evaluated module 2.5 documents followed the CTD structure for the level-1 headings; 2.5.1 Product development rationale, 2.5.2 Overview of biopharmaceuticals, 2.5.3 Overview of clinical pharmacology, 2.5.4 Overview of efficacy, 2.5.5 Overview of safety, and 2.5.6 Benefits and risks conclusions. Document E-1 did not include a 2.5.2 level-1 heading and document E-2 had a different order of the level-1 headings than what is recommended by the CTD. Document E-1 and E-2 were from the same sponsors and were written for antiviral and oncology indications, respectively. One document (G-1) had an additional level-1 heading, describing data in the context of historical clinical trials and

[10] Sponsor-specific is referred to as a trend that is specified for a specific sponsor and its two documents (document X-1 and X-2), meaning a trend that is seen in one sponsor might vary greatly from what is seen in another sponsor. Not sponsor-specific means, therefore, that a trend that is seen in one document from a specific sponsor, does not necessarily have to be observed in another document from the same sponsor

real-world data. All sponsors (100%) presented a reference list and three documents (15%) included extra data in the form of a separate appendix.

Table 15. Structure of module 2.5 documents according to the CTD standards

Document	Executive summary	2.5.1	2.5.2	2.5.3	2.5.4	2.5.5	2.5.6	Additional level-1 headings	Reference list	Appendix
A-1	-	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
A-2	-	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
B-1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	Provided seperately
B-2	-	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
C-1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
C-2	-	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
D-1	-	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
D-2	-	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
E-1	-	Yes	-	Yes	Yes	Yes	Yes	-	Yes	Provided seperately
E-2*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
F-1	-	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
F-2	-	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
G-1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	1	Yes	-
G-2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
H-1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
H-2	-	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
I-1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	Provided seperately
I-2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	
J-1	-	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-
J-2	-	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	-

- = Not mentioned; Yes = The section was provided; Provided separately = Not found in the same document
** the order of the level-1 headings is presented differently*

Table 16, which shows the structure of the module 2.7.1 documents, displays an executive summary in five (25%) module 2.7.1 documents. The majority (80%, 16/20) of the documents followed the CTD structure for the level-1 headings; 2.7.1.1 Background and overview, 2.7.1.2 Summary of results of individual studies, and 2.7.1.3 Comparison and analyses of results across studies. However, document E-1 did not include the level-1 heading 2.7.1.2. The same sponsor for document E-2 had a different structure compared to a typical 2.7.1 document, hence, the document could not be compared to the other documents. The level-1 heading 2.7.1.3 of document B-1 was worded differently (Analytical methods) than in the CTD guideline. Document A-2 included two additional level-1 headings in the document, describing extended-release claim, and *in vitro* and *in vivo* correlation. A reference list was included in nine documents (45%) and a conclusion was included in five documents (25%). Majority of the documents (65%, 13/20) included extra data in the form of an appendix that was either included in the document or as a separate file.

Table 16. Structure of module 2.7.1 documents according to the CTD standards

Document	Executive summary	2.7.1.1	2.7.1.2	2.7.1.3	Additional level-1 headings	Conclusion	Reference list	Appendix
A-1	-	Yes	Yes	Yes	-	-	-	Yes
A-2	-	Yes	Yes	Yes	2	-	-	Yes
B-1	-	Yes	Yes	Different title	-	-	Yes	-
B-2	-	Yes	Yes	Yes	-	-	Yes	Yes
C-1	Yes	Yes	Yes	Yes	-	-	-	Yes
C-2	-	Yes	Yes	Yes	-	-	-	-
D-1	-	Yes	Yes	Yes	-	-	-	Provided seperately
D-2	-	Yes	Yes	Yes	-	-	-	Provided seperately
E-1	-	Yes	-	Yes	-	-	Yes	Yes
E-2*	---------- Not following the same structure as a common 2.7.1 document ----------							
F-1	-	Yes	Yes	Yes	-	-	-	-
F-2	-	Yes	Yes	Yes	-	Yes	Yes	-
G-1	Yes	Yes	Yes	Yes	-	-	Yes	Yes
G-2	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes
H-1	-	Yes	Yes	Yes	-	Yes	Yes	Yes
H-2	---------- No document provided ----------							
I-1	Yes	Yes	Yes	Yes	-	-	-	Yes
I-2	Yes	Yes	Yes	Yes	-	-	-	Provided seperately
J-1	-	Yes	Yes	Yes	-	Yes	Yes	Yes
J-2	-	Yes	Yes	Yes	-	Yes	Yes	No info

- = Not mentioned; Yes = The section was provided; Provided separately = Not found in the same document; No info = The section was provided but no information was applicable or available; Different title = The section was mentioned with different words than in accordance to the CTD standard
** This document is not following the CTD structure as a standard 2.7.1 document*

Table 17 displays the structure of module 2.7.2 documents. An executive summary was included in four (20%) documents. All documents followed the CTD structure for the level-1 headings; 2.7.2.1 Background and overview, 2.7.2.2 Summary of results of individual studies, and 2.7.2.3 Comparison and analyses of results across studies, whilst the level-1 heading 2.7.2.4 Special studies was only included in 35% (7/20) of the documents. In 12 of the documents the level-1 heading 2.7.2.4 was either not included (15%, 3/20) or included with no information provided (45%, 9/20). In addition, the level-1 heading 2.7.2.4 for document E-1 was titled differently (virology summary) when compared to the CTD standard, and document E-2 was the only document that had included an additional level-1 heading 2.7.2.5, describing summary of key findings. A reference list was included in 12 module 2.7.2 documents (60%) and a conclusion in four documents (20%). Majority of the documents (70%, 14/20) included extra data in the form of an appendix that was either included in the document or as a separate file.

Table 17. Structure of module 2.7.2 documents according to the CTD standards

Document	Executive summary	2.7.2.1	2.7.2.2	2.7.2.3	2.7.2.4	Additional level-1 headings	Conclusion	Reference list	Appendix
A-1	-	Yes	Yes	Yes	Yes	-	-	-	No info
A-2	-	Yes	Yes	Yes	-	-	-	-	-
B-1	-	Yes	Yes	Yes	-	-	-	Yes	-
B-2	-	Yes	Yes	Yes	Yes	-	-	Yes	Yes
C-1	Yes	Yes	Yes	Yes	No info	-	-	-	Yes
C-2	-	Yes	Yes	Yes	No info	-	-	-	Yes
D-1	-	Yes	Yes	Yes	No info	-	-	-	Provided seperately
D-2	-	Yes	Yes	Yes	No info	-	-	-	Provided seperately
E-1	-	Yes	Yes	Yes	Different title - not CTD	-	-	Yes	Yes
E-2	-	Yes	Yes	Yes	Yes	1	Discussion	Yes	-
F-1	-	Yes	Yes	Yes	Yes	-	Yes	-	-
F-2	-	Yes	Yes	Yes	Yes	-	-	Yes	No info
G-1	Yes	Yes	Yes	Yes	Yes	-	-	Yes	Yes
G-2	Yes	Yes	Yes	Yes	-	-	-	Yes	Yes
H-1	-	Yes	Yes	Yes	Yes	-	-	Yes	Yes
H-2	-	Yes	Yes	Yes	No info	-	-	Yes	Yes
I-1	-	Yes	Yes	Yes	No info	-	Yes	Yes	Yes
I-2	-	Yes	Yes	Yes	No info	-	Yes	Yes	Yes
J-1	-	Yes	Yes	Yes	No info	-	-	Yes	Yes
J-2	-	Yes	Yes	Yes	No info	-	-	-	Yes

- = Not mentioned; Yes = The section was provided; Provided separately = Not found in the same document; No info = The section was provided but no information was applicable or available; Different title - not CTD = The section was mentioned with different information than in accordance to the CTD standard; Discussion = Different naming than conclusion

Table 18 shows the structure of module 2.7.3 documents. An executive summary was included in six (30%) documents. All documents followed the CTD structure for the level-1 headings; 2.7.3.1 Background and overview of clinical efficacy, 2.7.3.2 Summary of results of individual studies, 2.7.3.3 Comparison and analyses of results across studies, 2.7.3.4 Analysis of clinical information relevant to dosing recommendations, and 2.7.3.5 Persistence of efficacy and/or tolerance effects. Two documents (10%, document F-2 and G-1) worded the level-1 heading 2.7.3.2 differently (Summary of results for study X and Summary of results of individual drug Y monotherapy studies) when compared to the CTD guideline. Three documents (15%) included one or several additional level-1 headings, describing efficacy conclusions, drug combination studies, studies of specific diseases and single patient protocols. A reference list and conclusion were included in 60% (12/20) and 35% (7/20) of the documents, respectively. Majority of the documents (75%, 15/20) included extra data in the form of an appendix that was either included in the document or as a separate file.

Table 18. Structure of module 2.7.3 documents according to the CTD standards

Document	Excecutive summary	2.7.3.1	2.7.3.2	2.7.3.3	2.7.3.4	2.7.3.5	Additional level-1 headings	Conclusion	Reference list	Appendix
A-1	-	Yes	Yes	Yes	Yes	Yes	-	-	-	Yes
A-2	Yes	Yes	Yes	Yes	Yes	Yes	-	-	-	Yes
B-1	-	Yes	Yes	Yes	Yes	Yes	-	-	Yes	-
B-2	-	Yes	Yes	Yes	Yes	Yes	-	-	Yes	Yes
C-1	Yes	Yes	Yes	Yes	Yes	Yes	-	-	-	-
C-2	-	Yes	Yes	Yes	Yes	Yes	-	-	-	Yes
D-1	-	Yes	Yes	Yes	Yes	Yes	-	-	-	Provided seperately
D-2	Yes	Yes	Yes	Yes	Yes	Yes	-	-	-	Provided seperately
E-1	-	Yes	Yes	Yes	Yes	Yes	-	-	Yes	Yes
E-2	-	Yes	Yes	Yes	Yes	Yes	1	Yes	Yes	Yes
F-1	-	Yes	Yes	Yes	Yes	Yes	-	Yes	-	-
F-2	-	Yes	Different title	Yes	Yes	Yes	-	Yes	Yes	Yes
G-1	-	Yes	Different title	Yes	Yes	Yes	1	Yes	Yes	Yes
G-2	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	Yes	-
H-1	-	Yes	Yes	Yes	Yes	Yes	-	-	Yes	Yes
H-2	-	Yes	Yes	Yes	Yes	Yes	-	-	Yes	Yes
I-1	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes
I-2	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	Yes	-
J-1	-	Yes	Yes	Yes	Yes	Yes	-	-	Yes	Yes
J-2	-	Yes	Yes	Yes	Yes	Yes	3	-	-	Yes

- = Not mentioned; Yes = The section was provided; Provided separately = Not found in the same document; Different title = The section was mentioned with different words than in accordance to CTD standard

Table 19 shows the structure of module 2.7.4 documents. An executive summary was included in eight (40%) documents. Sixteen documents (80%) followed the CTD structure for the level-1 headings; 2.7.4.1 Exposure to the drug, 2.7.4.2 Adverse events, 2.7.4.3 Clinical laboratory evaluations, 2.7.4.4 Vital signs, physical findings, and other observations related to safety, and 2.7.4.5 Safety in special groups and situations, whilst only 5 documents (25%) included information for the level-1 heading 2.7.4.6 Post-marketing data. Document F-2 mentioned the level-1 heading 2.7.4.2 as a subheading of the level-1 heading 2.7.4.1, and level-1 headings 2.7.4.3 and 2.7.4.4 were mentioned as subheadings of an additional level-1 heading (not in accordance with the CTD), therefore, only two level-1 headings from the CTD guideline were referred to in the F-2 document. A similar exception was seen for document C-2, where the level-1 heading 2.7.4.6 was mentioned as a subheading of level-1 heading 2.7.4.5. Document I-2 worded the level-1 heading 2.7.4.1 differently (Exposure to drug Z) than the CTD. Document D-2 could not be compared to the other documents since it included several studies and had an alternative structure. Two documents (10%) included additional level-1 headings, describing background and overview, summaries, subject disposition, study population demographics and baseline characteristics, correlative analysis pharmacokinetic and pharmacodynamic outcomes, additional safety studies and narratives. A reference list and conclusion were included in 60% (12/20) and 35% (7/20) of the documents, respectively. Majority of the documents (70%, 14/20) included extra data in the form of an appendix that was either included in the document or as a separate file.

Table 19. Structure of module 2.7.4 documents according to the CTD standards

Document	Executive summary	2.7.4.1	2.7.4.2	2.7.4.3	2.7.4.4	2.7.4.5	2.7.4.6	Additional level-1 headings	Conclusion	Reference list	Appendix
A-1	-	Yes	Yes	Yes	Yes	Yes	No info	-	-	-	Yes
A-2	Yes	Yes	Yes	Yes	Yes	Yes	No info	-	-	-	Yes
B-1	-	Yes	Yes	Yes	Yes	Yes	Yes	-	-	Yes	Provided seperately
B-2	-	Yes	Yes	Yes	Yes	Yes	No info	-	-	Yes	Yes
C-1	Yes	Yes	Yes	Yes	Yes	Yes	No info	-	-	-	-
C-2	-	Yes	Yes	Yes	Yes	Yes	Mentioned as a subheading	-	-	-	Yes
D-1	Yes	Yes	Yes	Yes	Yes	Yes	No info	-	-	-	Provided seperately
D-2	----------This document includes many different studies and cannot be compared with the other documents----------										
E-1	-	Yes	Yes	Yes	Yes	Yes	No info	-	-	Yes	Yes
E-2	-	Yes	Yes	Yes	Yes	Yes	Yes	7	Yes	Yes	-
F-1	-	Yes	Yes	No info	Yes	Yes	Yes	-	-	-	-
F-2*	-	Yes	Mentioned as a subheading	Mentioned as a subheading	Mentioned as a subheading	Yes	No info	3	Yes	Yes	Yes
G-1	Yes	Yes	Yes	Yes	Yes	Yes	No info	-	Yes	Yes	-
G-2	Yes	Yes	Yes	Yes	Yes	Yes	No info	-	Yes	Yes	-
H-1	-	Yes	Yes	Yes	Yes	Yes	Yes	-	-	Yes	Yes
H-2	-	Yes	Yes	Yes	Yes	Yes	No info	-	-	Yes	Yes
I-1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes + Provided seperately
I-2	Yes	Different title	Yes	Yes	Yes	Yes	No info	-	Yes	Yes	Yes + Provided seperately
J-1	Yes	Yes	Yes	Yes	Yes	Yes	No info	-	Yes	Yes	Yes
J-2	-	Yes	Yes	Yes	Yes	Yes	No info	-	-	-	Yes

- = Not mentioned; Yes = The section was provided; Provided separately = Not found in the same document; No info = The section was provided but no information was applicable or available; Different title = The section was mentioned with different words than in accordance to the CTD standard;
* section 2.7.4.2, 2.7.4.3, 2.7.4.4 are mentioned as subheadings and are included in other level-1 headings in the document. Only two level-1 headings of the CTD guidelines were mentioned in the document according to the standards.

5.5.1 What were the structural differences in the 20 module 2 packages from the different sponsors?

The level-1 headings had multiple variations, however, no document was incorrectly structured as the guidance is only a recommendation. This is another example that shows that recommendations will create differences in the documentation process between different sponsors.

An executive summary was included in 31% (31/100) of the documents, but this was not a sponsor-specific action, meaning that a sponsor that had included the executive summary in one document may not have included it in all other documents. Module 2.5 clinical overview and module 2.7.4 safety summary documents had included an executive summary more often (40% for both module 2.5 and 2.7.4) than other documents (25% for module 2.7.1, 20% for module 2.7.2 and 30% for module 2.7.3). The executive summary was described as a short overview of the document, though, not always mentioned with the specific wording "executive summary", but could be mentioned as "summary of clinical pharmacology/summary of biopharmaceutic studies and analytical methods/ summary of clinical pharmacology studies" and so forth, depending on the document module.

The numbering of the level-1 headings differed between sponsors and, for example, the first appearing level-1 heading was either numbered as "2.5.1" or as "1", depending on the sponsor. Sponsors B, C, E, F, G, H and J numbered according to the latter. The approach chosen was sponsor-specific, meaning all module 2 documents from one sponsor used the same numbering style.

In one document (1%, 1/100), the order of the level-1 headings was presented differently and in five (5%, 5/100) of the documents the heading wording was not identical to the CTD wording but was modified to fit the presented content.

Altogether, five sponsors included one or multiple additional level-1 headings. Document A-2 included two extra level-1 headings in module 2.7.2, document F-2 included three extra level-1 headings in module 2.7.4 and document J-2 included three extra level-1 headings in module 2.7.3. Document G-1 included one extra level-1 heading in module 2.5 and one in module 2.7.3. Sponsor E-2 included one extra level-1 heading in module 2.7.1 and module 2.7.3, and seven extra level-1 headings in module 2.7.4. The inclusion of additional level-1 headings was not a sponsor-specific action. The added level-1 headings were not necessarily included in the end of the document but in-between the standard CTD level-1 headings, which resulted in numbering changes of the headings. The additional level-1 headings did not have a relationship with the total page count of the specific document, i.e., it did not indicate that the total number of pages was higher than average or closer to the highest page count for the documents where the additional level-1 headings were found.

If no data were found for a specific level-1 heading, the section was either mentioned and then followed by the text "not applicable" or "no available information", or completely removed from the document. These were sponsor-specific approaches.

The reference list was included in the end of the module 2.5 documents for all sponsors. For the 2.7 summary documents (modules 2.7.1, 2.7.2, 2.7.3, 2.7.4), the reference list was not always included as a separate level-1 heading but included in the appendix as a subheading. The inclusion and location of a reference list in a specific document was sponsor-specific for 9 out of 10 sponsors. Sponsors B, E, G, H, I, J and document F-1 included reference lists in at least one of the module 2.7 documents (2.7.1, 2.7.2, 2.7.3, 2.7.4). Sponsors A, C, D and document F-2 did not include a reference list in any of the module 2.7 documents.

Documents D-2, E-2 and H-2 had specific exceptions for certain modules. In module 2.7.1, the E-2 document did not follow the same structure as the other module 2.7.1 documents, while document H-2 did not provide any document for the same module. In module 2.7.4, document D-2 included many different studies that made a comparison with the other documents impossible. These exceptions are further reflected in the subheading figures (section 5.6).

5.5.2 How well is the CTD guideline followed?

An executive summary provides a short introduction to the rest of the document, however, it is not according to the CTD standard and does not have to be included in the documents.

The CTD guideline provides recommendations on how to structure, number and name the level-1 headings in the table of contents. The following structure was recommended for numbering the level-1 headings: 2.5.1, 2.5.2, 2.5.3…" for clinical overview and "2.7.1.1, 2.7.2.2 …" for summary documents, and so forth. The recommendation was followed by three (30%, 3/10) sponsors. The other sponsors used a single digit numbering system "1, 2, 3, 4…" for all documents. The deviation from the CTD structure is not incorrect and the heading numbering as 1, 2, 3 … follows simple logic and was accepted by the health authorities. Arabic numbers are to be used in accordance with the CTD guideline and this recommendation was followed by all sponsors for both the level-1 heading and subheading numbering. The CTD recommended structure was followed by 99% (99/100) of the documents.

The CTD guideline states that one reference list should be provided for module 2.5 documents and one should be included for the total clinical summary document 2.7 (meaning all module 2.7.1, 2.7.2, 2.7.3, 2.7.4 combined). The reference list for the total clinical summary document 2.7 can also be presented as a separate section (2.7.5) after the total clinical summary document, meaning that no reference list must be provided within the summary documents if a separate section is provided. The reference lists were only analyzed for those documents that had provided the references within the individual documents. A reference list provided as a separate section (2.7.5) was not analyzed. As sponsors A, C, D and document F-2 did not provide any reference list in any of the summary documents (module 2.7) it can be assumed that a separate reference list (2.7.5) was provided. A variation was shown in how reference lists were included for the different clinical summary documents. While one reference list was mentioned in one document, it was left out in another, i.e., a sponsor did not necessarily follow the same approach for all summary documents.

Conclusions were sometimes included as a level-1 heading, however, nothing is specified about including a conclusion as a level-1 heading in the CTD guideline. The only specific level-1 heading mentioning a conclusion is the "benefit and risk conclusion" heading in module 2.5.

According to CTD, an appendix should be included at the end of a document if there are any detailed presentations of methods and results. Lengthy tables should also be provided in the appendix. An appendix should be placed after a specific module (2.5, 2.7.1, 2.7.2 …). It is also specified in module 2.5 section 2.5.6 that an additional appendix (section 2.5.6.5) can be included if a detailed presentation of the summarized section 2.5.6.4 (benefit-risk assessment) is necessary. Some sponsors have also included an appendix as a separate document, an approach that is not mentioned in the CTD recommendations.

5.6 Subheadings

Figure 16 to Figure 20 were created to show the total number of subheadings in relation to the number of pages in the text body. Only the subheadings that belong to the level-1 headings, as established in the level-1 heading tables above, were calculated. In addition, the number of pages that belong to the level-1 headings as per the CTD structure has been calculated to form a subheading versus page count ratio. The level-1 headings according to the CTD structure are those presented in section 2.1, Table 2. This allows the analysis of how much text is provided per subheading and how detailed the documents are.

The presentation of this data will merely focus on the subheading to page count ratio, and not on the document length as this has been brought up in section 5.2.

In figures 16 to 20, the y-axis shows both the number of subheadings and the number of pages. The x-axis presents the specific sponsors and documents. The blue bars show the total number of subheadings for the level-1 headings in the document, and the orange bars show the total page count for the level-1 headings.

The ratio between the blue and orange bars are interpreted so that when the number of pages is higher than the number of subheadings it implies that one subheading has a content length over one page. Similarly, a higher number of subheadings compared to the number of pages indicates that there is less than one page of content per subheading.

Figure 16 shows the module 2.5 documents. It was observed that the sponsors are to some extent aligned when investigating the relationship between subheadings and the number of pages. The number of subheadings was close to the number of pages for 40% (8/20) of the sponsors. Sponsor A showed greater difference in subheading versus number of pages, ratio 2.8 and 2.1 for document A-1 and A-2, respectively (meaning that 2.8 and 2.1 number of pages was the average content amount per subheading), compared to other sponsors.

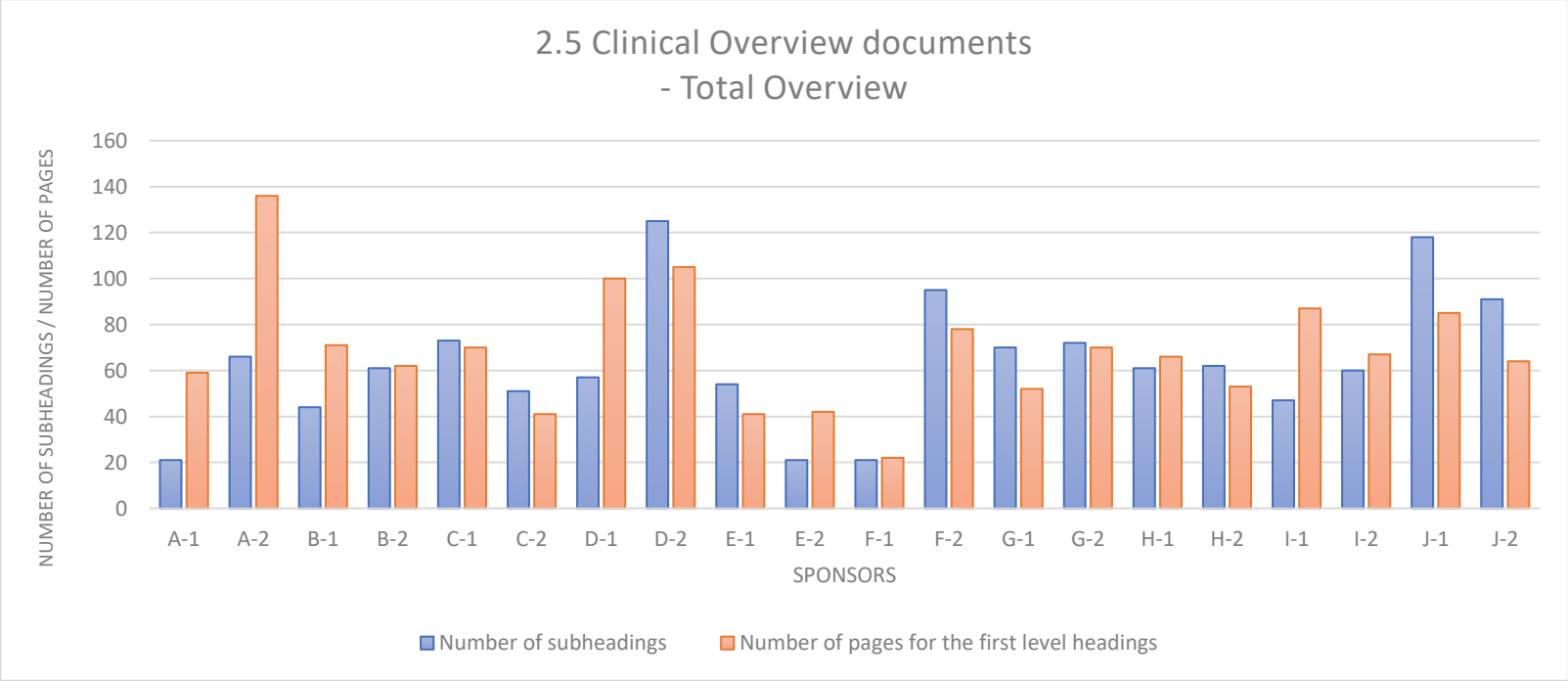

Figure 16. Number of subheadings and number of pages in module 2.5.

Figure 17 shows that a clear trend is followed in the module 2.7.1 documents. For the majority (70%, 14/20) of the documents, it was distinguished that the number of pages were expressed in higher quantity than the number of subheadings. This trend was module-specific[11] rather than sponsor-specific. Document I-1 had a 3.3 ratio and contained the highest number of pages. Document C-1 and G-2 had more subheadings than number of pages and both presented a 0.8 ratio. No subheading data were available for document E-2 as the document followed an unusual 2.7.1 structure. In addition, document H-2 did not contain a 2.7.1 document.

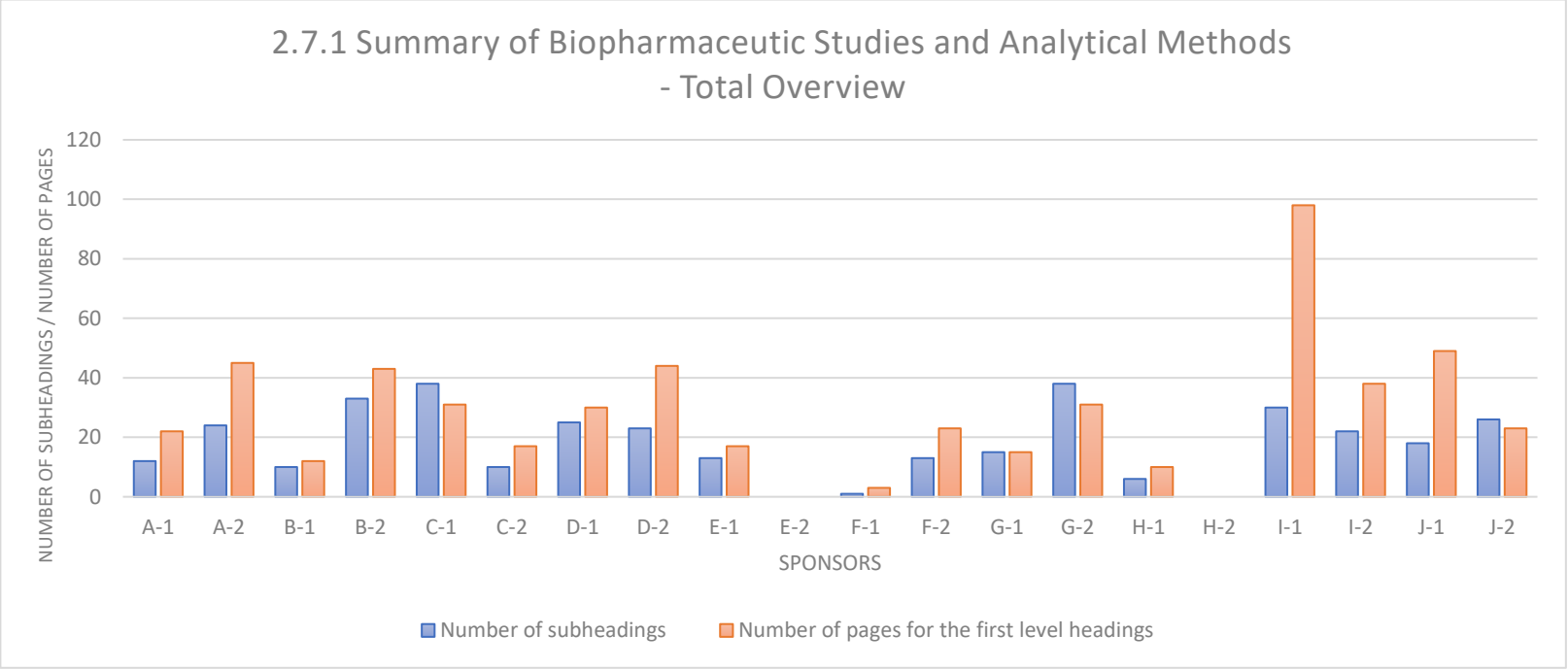

Figure 17. Number of subheadings and number of pages in module 2.7.1.

[11] Module-specific is referred to as a trend that is specified for a specific module, meaning a trend that is seen in one module should be followed to a great extent by another sponsor in the same module

Figure 18 shows that module 2.7.2 documents followed the same trend as in Figure 17, but the number of pages varied more between the sponsors than the number of subheadings. This trend also appeared to be module-specific rather than sponsor-specific. Sponsor G (ratio 0.8) and document C-1 (ratio 0.9) were clear exceptions in that and, in this occasion, they appeared to favor a greater number of subheadings in relation to the number of pages. Some documents (e.g., E-1, I-1, J-1) had a ratio 5 between the number of pages and number of subheadings, i.e., reaching up to 5 pages of content per subheading.

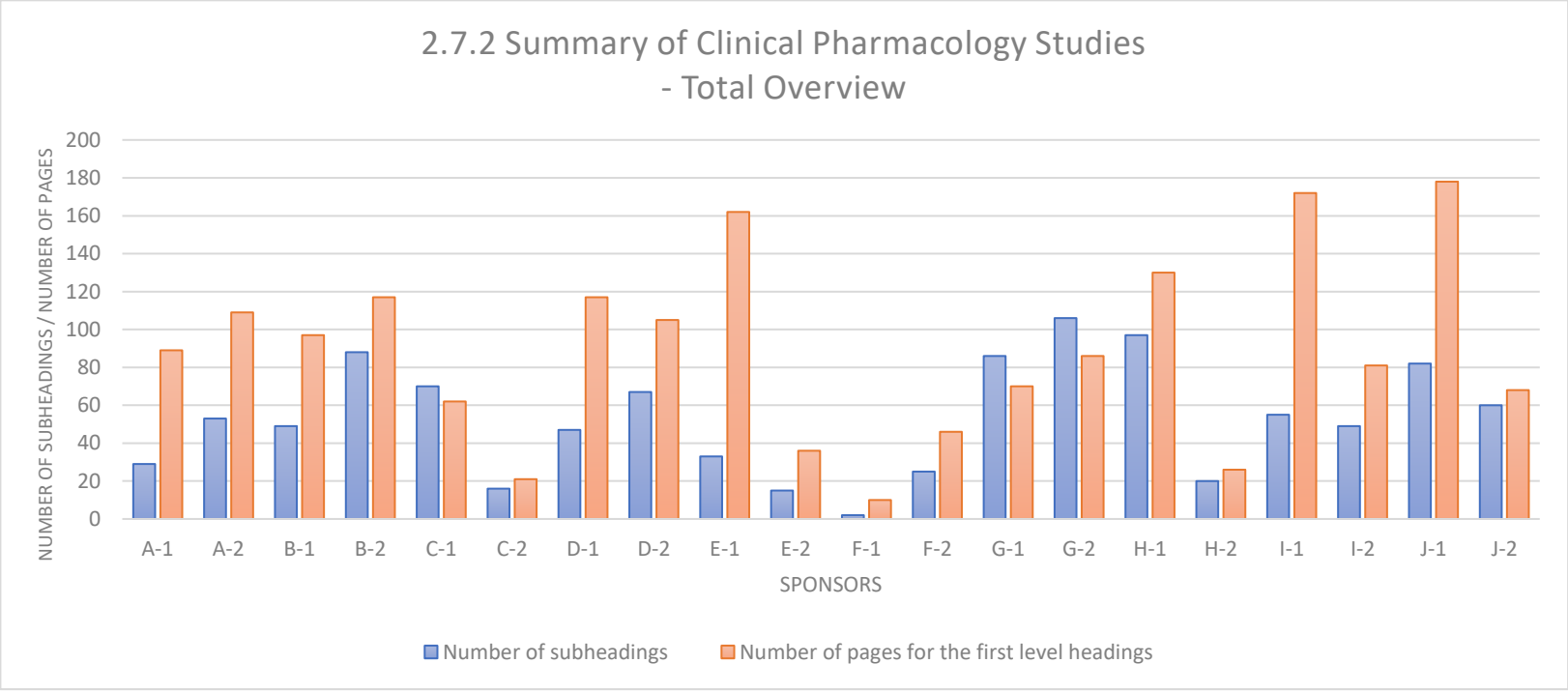

Figure 18. Number of subheadings and number of pages in module 2.7.2.

Figure 19 shows that the module 2.7.3 documents followed the same trend as shown in Figure 17 and Figure 18. For most documents (85%, 17/20), the total number of pages was expressed in a higher quantity than the number of subheadings. Document A-2 had the largest ratio (5.8) difference between the number of pages and number of subheadings. Visible exceptions were observed continuously for sponsor G (ratio 0.8 and 0.9, respectively) and also document H-2 (ratio 0.8).

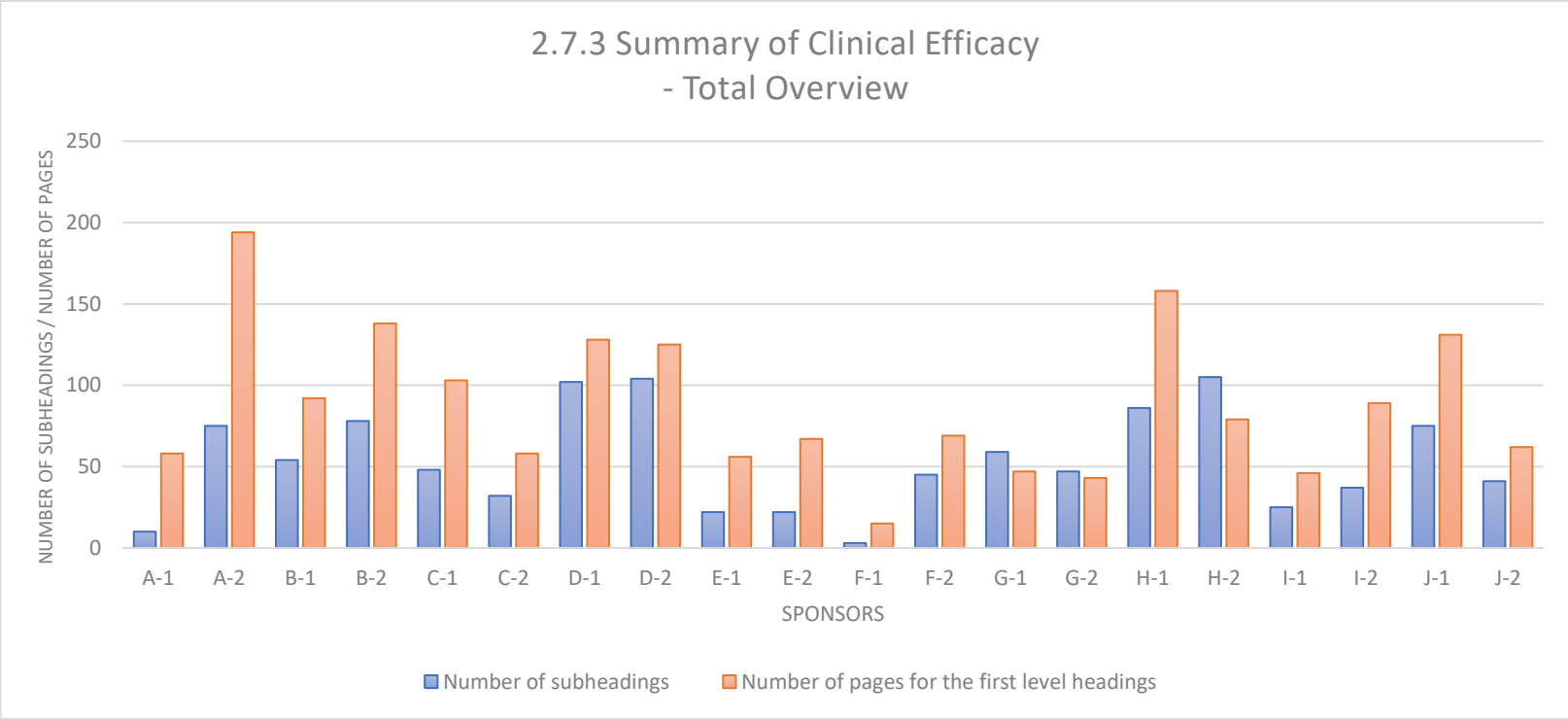

Figure 19. Number of subheadings and number of pages in module 2.7.3.

In Figure 20, it was observed that half of the sponsors (50%, 5/10) for the module 2.7.4 documents followed a sponsor-specific approach when investigating the ratio between subheadings and the number of pages. For the majority of documents (75%, 15/20), the total number of pages is displayed in a higher granularity than the number of subheadings, however, this was not the trend for sponsor G (ratio 0.7 and 1.1) and document F-1 (ratio 0.6) and H-2 (ratio 0.8). Document A-2 continued to have the largest ratio (3.4) difference between the number of pages and number of subheadings. No data were compiled for document D-2 since the document contained many different studies and could not be compared to the other documents.

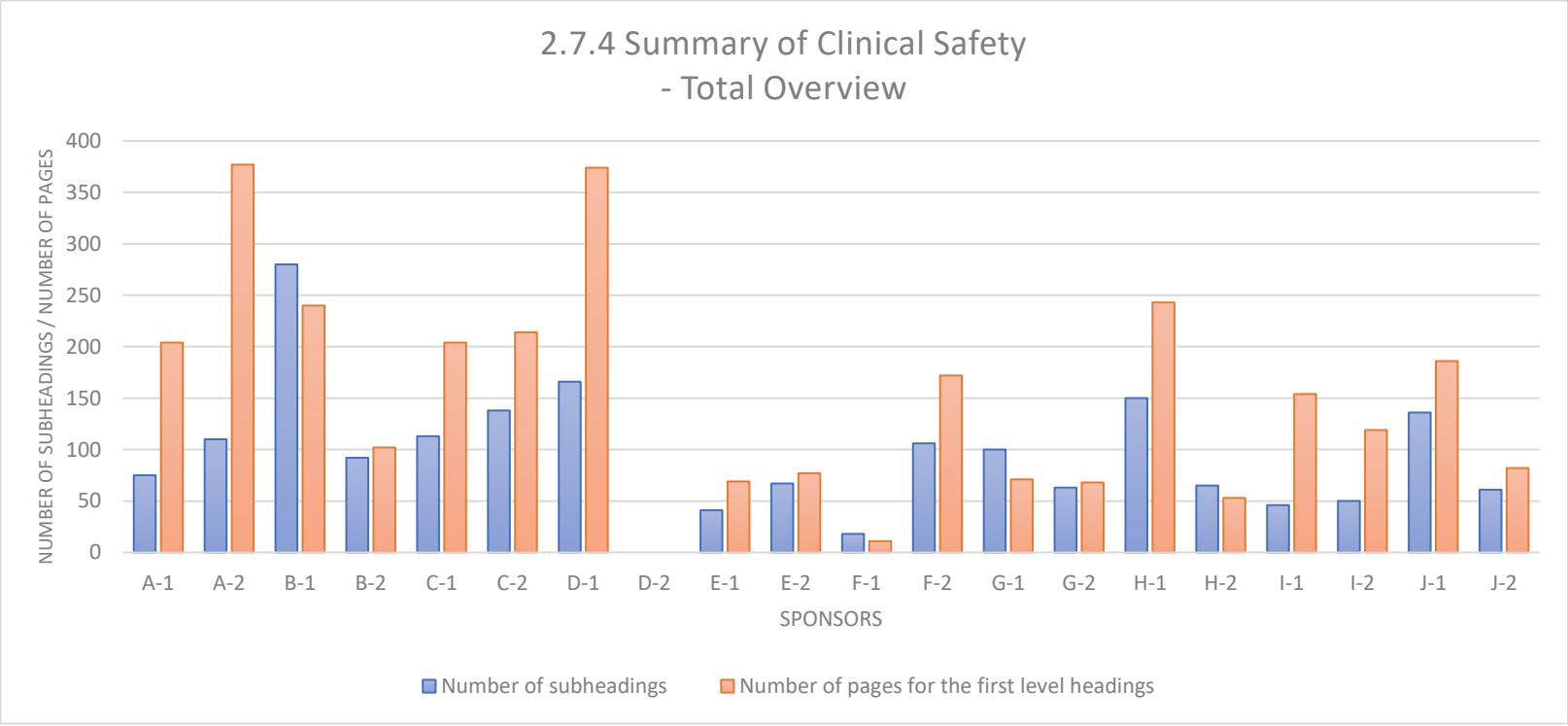

Figure 20. Number of subheadings and number of pages in module 2.7.4.

5.6.1 What were the structural differences in the 20 module 2 packages from the different sponsors?

The different modules tended towards higher number of pages in relation to the number of subheadings, i.e., one subheading had a content length over one page. An exception was module 2.5 where the number of pages and number of subheadings was near equal, i.e., one page of content per subheading. Some trends were aligned through all sponsors (e.g., the quantity of subheadings and page count) and some tendencies appeared to be sponsor-specific (e.g., the ratio between these variables).

Sponsor G appeared to slightly favor a greater number of subheadings in relation to the number of pages for all documents regardless of the module, which indicates documents with short subsections and several subheadings per page. Sponsor A had long documents structured with few subheadings and this was true for both document A-1 and A-2. Other sponsors had a tendency to include more subheadings which created documents with shorter content description per subheading. Sponsor I followed the same trend but the difference between number of pages and number of subheadings was not as significant for all modules as was shown for sponsor A.

5.6.2 How well is the CTD guideline followed?

The results showed that sponsors tend to structure the content by including subsections in a higher number than what is mentioned in the CTD guideline.

The module 2.5 document has only seven subheadings mentioned in the CTD structure within the level-1 heading 2.5.6 (2.5.6.1, 2.5.6.1.1, 2.5.6.1.2, 2.5.6.2, 2.5.6.3, 2.5.6.4, 2.5.6.5). According to Figure 16, all documents (100%) contained 21 or more subheadings.

The CTD does not include recommendations for subheading grouping for module 2.7.1 or 2.7.2, but the majority of the module 2.7.1 documents (80%) and module 2.7.2 documents (95%) contained 10 or more subheadings. However, there are other recommendations based on the number of summarized studies and how these should be structured, which will result in more content.

The module 2.7.3 documents have three subheadings mentioned in the CTD structure within the level-1 heading 2.7.3.3 (2.7.3.3.1, 2.7.3.3.2, 2.7.3.3.3). According to Figure 19, one document (F-1) included three subheadings, whilst the rest of the documents included 10 or more subheadings. The greatest number of subheadings was 105 in document H-2. In other words, only one document (F-1) had followed the guidance on how many subheadings to include in module 2.7.3 documents.

The level-1 heading 2.7.4.1 for the module 2.7.4 document is recommended to be structured in three subheadings (2.7.4.1.1, 2.7.4.1.2, 2.7.4.1.3) and level-1 heading 2.7.4.2 is recommended to be structured further by two subheadings (2.7.4.2.1, 2.7.4.2.2). In addition, level-1 heading 2.7.4.5 is recommended to be structured further by eight subheadings (2.7.4.5.1, 2.7.4.5.2, 2.7.4.5.3, 2.7.4.5.4, 2.7.4.5.5, 2.7.4.5.6, 2.7.4.5.7, 2.7.4.5.8). In other words, 13 subheadings in total are recommended to be included in the module 2.7.4 document. All (100%, 19/19) of the documents had included more than 13 subheadings in module 2.7.4, hence, no sponsor followed the recommended guidance. Document D-2 was not included in the calculations, as the document included several studies and could, therefore, not be compared to the other documents.

5.7 Repetition of table data in the text

Figure 21 and Figure 22 show the repetition of table content in the text across different sponsors in module 2.5 and 2.7.4 documents. The left y-axis of the figures presents the number of tables for each document and the right y-axis presents the number of pages. The x-axis presents a specific sponsor and document. Note the scale variance of the y-axis for the respective tables.

The data are based on four colored categories:

Category 1. Data from table is not repeated in text

Category 2. Data from table is summarized in text

Category 3. Data from table is somewhat repeated in text

Category 4. Data from table is (almost) completely repeated in text

Please visit section 4.3 and Table 9 for a detailed explanation of the categories.

Figure 21 shows a sponsor-specific variation between the data repetition categories in the reviewed module 2.5 documents. In all documents, table data are repeated in the text to some extent. Category 1, 2 and 3 repetition was seen in all 20 module 2.5 documents. In 15 of the documents, more than 75% (15/20) of all tables in each document fell under category 1 or 2. The rest of the tables in the module 2.5 documents fell under category 3, except for six documents (30%) that had one or two tables in category 4. Document C-2, D-2, E-1, E-2 and I-1 had one table in category 4, while document F-2 had two tables. The use of category 4 was document-specific rather than sponsor-specific. Half of the documents that presented category 4 tables (3/6, 50%) had 100 or more pages in their documents.

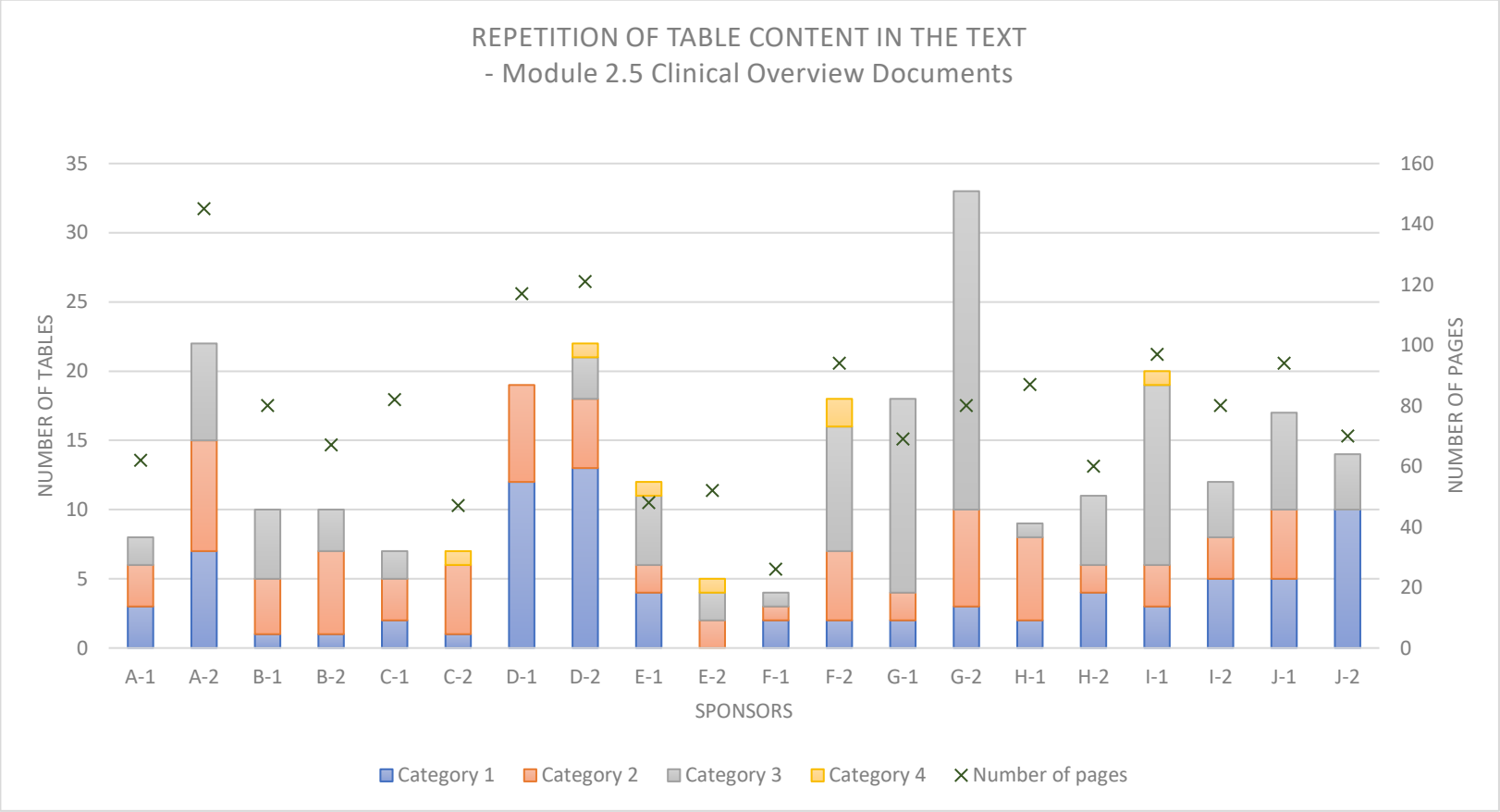

Figure 21. Module 2.5 document variation for repetition of table data in the text.

Figure 22 shows the table data repetition levels in the module 2.7.4 documents. The most frequently seen category was 3, which was present in all documents, except document F-1, and appeared as the average category for 80% (16/20) of all the documents. Category 2 was present in 90% (18/20) of the documents but only as the average category for 10% (2/20). Category 1 was applied in 95% (19/20) of the documents but was only the average category for 10% of all the documents. Category 4 was applied in one or two tables for 45% (9/20) of the documents. Documents A-1, E-1, F-2, I-1 and J-1 had one table in category 4, while documents C-1, D-2 and E-2 had two tables. Document A-2 and D-1 presented documents with higher number of pages. A-2 presented all four table categories whilst the majority (61/77, 79%) of the D-1 tables where in category 1 and 2. In other words, table with less repetition did not indicate a shorter document.

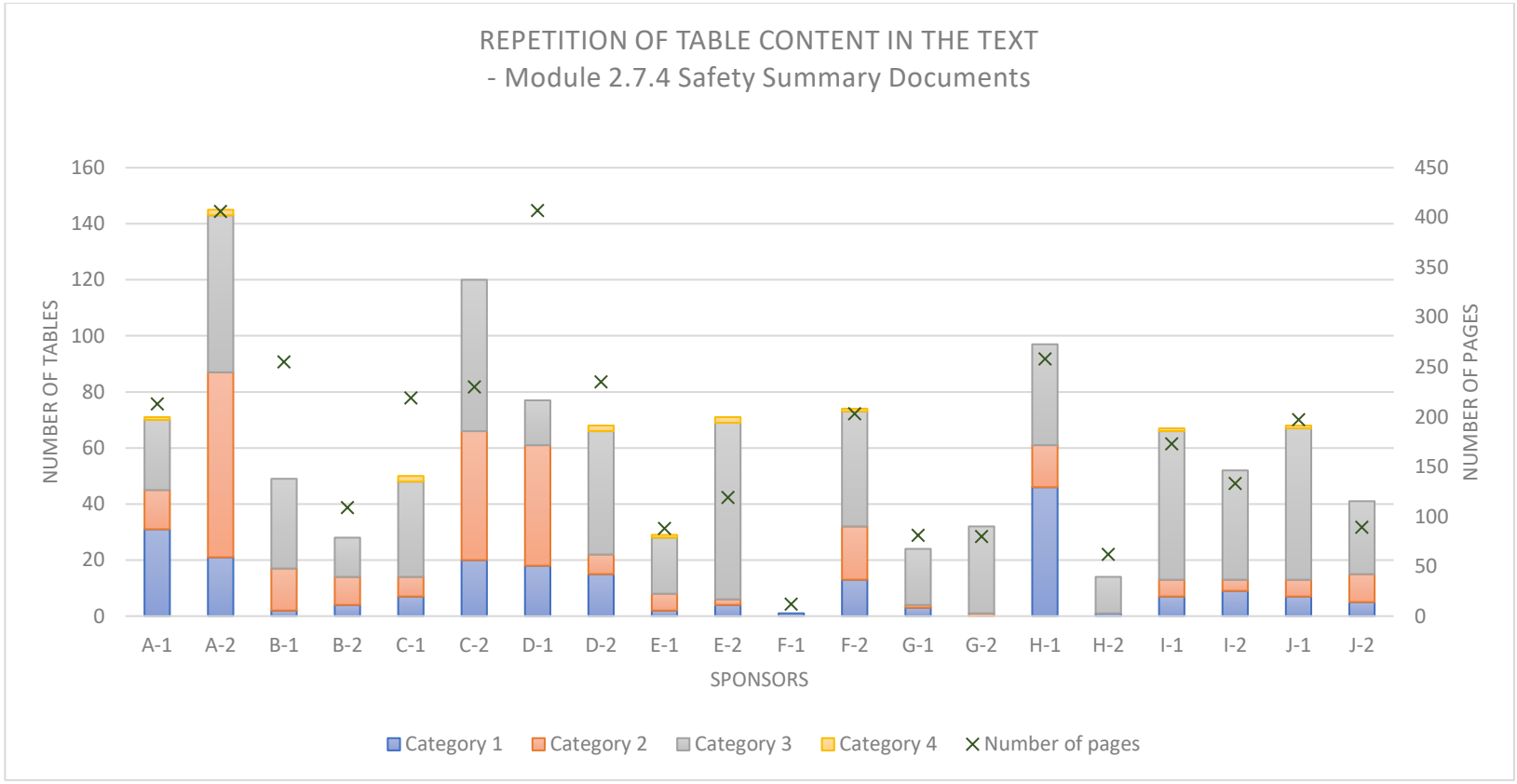

Figure 22. Module 2.7.4 document variation for repetition of table data in the text.

5.7.1 What were the structural differences in the 20 module 2 packages from the different sponsors?

Repetition occurred in all documents for module 2.5 and 2.7.4, and variation existed between sponsors regarding the level of repetition of table content in the text. It was visible that a few sponsors were aiming to prepare their documents with less table content repetition, however, the differences were minor and could depend on unknown factors. For example, sponsor D aimed for less repetition in both documents for module 2.5, one document being for an oncology indication and the other for an RA indication. In module 2.7.4, the trend was not as evident for sponsor D, compared to the other sponsors.

Category 1 was often presented in a higher ratio in module 2.5 compared to module 2.7.4 documents. Category 2 was favored in module 2.5, while category 3 was preferred in module 2.7.4 documents. Category 4 was not considered as an ideal approach by any sponsor in the evaluated modules. Individual differences were identified within a sponsor, meaning a category ratio for one sponsor in one document did not necessarily reflect the same category ratio for another document of the same sponsor.

Both the number of tables and the table content repetitions had an impact on the document length, as expected. The correlation between the document length and number of tables is described in section 5.3.1. It is mentioned that the document length increased with a growing number of tables. The correlation between a table category and the document length was similar to the correlation between the document length and number of tables, but there were also documents (D-1 and A-2) that were long although the majority of tables were obtained in the lower repetition categories (category 1 and 2).

The research project did not measure the repetition of plain text between different document sections or any other type of repetition that might occur in a document. This kind of repetition might have a great impact on the document length.

The author leaves room for misinterpretation of the tables, as they were categorized personally by the author, with one pair of eyes alone, and wants to highlight the importance of this approach only being a base to understand differences between sponsors and documents with regards to repetition of table data in the text.

5.7.2 How well is the CTD guideline followed?

According to the CTD guideline, the module 2.5 clinical overview should not contain repetition of content that is fully presented elsewhere. It should also be a short document including tables and graphs to aid understanding and to establish rapid reading. There are no specific recommendations on how to repeat data from tables in the text. Module 2.7.4 documents contain more detailed descriptions and more repetition of table data in the text is expected, than for module 2.5 documents. This was also true, as shown by the collected data from the module 2.5 and 2.7.4 documents. Cross-referencing from module 2.5 to module 2.7 and module 5 would be beneficial to avoid unnecessary repetition. Similarly, would cross-referencing from module 2.7.4 to module 5 be advantageous.

5.8 Summary

What kind of documentation simplification could be done for module 2 documents?		
Clarification Variables	CTD standards are followed	Simplification suggestion
Number of pages	No	✓ Reduce document length in accordance with the CTD ✓ Reduce repetition of paragraphs and conclusions ✓ Number of pages can be reduced by following the below suggestions
Number of tables	Not applicable	✓ Generate only essential tables ✓ Cross-referencing where applicable
Number of figures	Not applicable	✓ Generate only essential figures ✓ Cross-referencing where applicable
Level-1 headings	Yes	✓ Consider whether it is vital to include extra level-1 headings that are not recommended by CTD
Subheadings	Yes[12]	✓ Consider whether it is vital to include extra subheadings that are not recommended by CTD
Table repetition	Not applicable	✓ Consider what has been presented in the tables and what is presented in the body text ✓ Reduce repetition ✓ Include cross-referencing

Not applicable = no specific CTD recommendation was mentioned for the variable

[12] The CTD recommendation was followed, but additional subheadings were included in a large number

6 Discussion

Today, when a drug is granted marketing authorization, it is important to remember that the documentation process of the drug does not end there. For example, the marketing authorization is followed by a redaction and anonymization process of certain MAA documents that will be made publicly available. Information will be exposed to the public and the parts that are to be published will continuously change and grow during the upcoming years. Therefore, it is essential to understand how to structure the documents and what information to include to ensure that both the authorities and the public are satisfied. For example, confidential information and individual patient data must stay protected throughout the whole process.

The strategy of presenting information in a report focuses mainly on the clinical trial endpoints, which will determine whether the drug is beneficial enough to be used in real life. This is the vital information that in the end will approve or reject a drug's marketing authorization. If we go back 10 years in time, the main focus of the process was to prove the success of the drug by including as much information as possible, a tactic that now must be restructured. The process has become more transparent and it is important that the sponsor creates writing strategies that take these changes into consideration. Instead of only focusing on proving the benefit of a drug, the writing strategy should start by focusing on what is necessary to grant a market approval overall, i.e., how much content is needed for a successful marketing authorization.

6.1 The CTD guideline

In general, the CTD guideline was well followed by all the sponsors. The level-1 headings and subheadings mentioned in CTD were all included in the documents. The inclusion of tables and figures were also followed, and the CTD guideline has limited description for how these should be used. The one CTD recommendation that was not followed was the document length. Therefore, the simplifications suggestions (see section 6.3) are examples of possible ways to shorten the document.

The CTD only provides recommendations for the sponsors on how to write and structure their documents. All recommendations emphasize the importance of a clear and easy-to-read document, which is a vague statement that leads to document variations and variation in the amount of presented information between different sponsors. The understanding of what a clear and easy-to-read document is varies between sponsors, or even between different study teams as part of the same sponsor. The data collected in this thesis show that writing strategies differs between sponsors and that the CTD recommendations are interpreted differently.

In general, it was expected that the total clinical 2.7 document would be larger than the module 2.5 clinical overview document, which was true for the collected results. What was not expected was that 95% of the module 2.5 documents had reached a total page count that was exceeding the 30-page CTD recommendation. Moreover, six documents had a total number of pages that measured threefold the CTD recommendation, i.e., 90 or more pages. For the total clinical 2.7 documents, 45% exceeded the 400-page CTD recommendation. In addition, one 2.7 document reached almost 800 pages, meaning the number of pages was almost twofold to what CTD recommends.

According to the CTD, module 2.7.1 should contain biopharmaceutic studies and associated analytical methods. This is the shortest document when the CTD content is compared to other module 2.7 documents. This was also true for the collected results. In accordance with the CTD, module 2.7.1 documents are only recommended to contain 3 level-1 headings and no subheadings (see section 5.5 and 5.6). A recommendation that was not followed. In addition, according to the CTD, it is expected that this document would have the fewest number of tables, which was true for the collected data.

Module 2.7.4 documents contain drug safety data and it was expected not only to be the longest document but also to provide the greatest number of tables, compared to other module 2.7 documents. This was true for the collected data. The CTD recommendation contains 6 level-1 headings and 15 subheadings for module 2.7.4. The safety in module 2.7.4 document is described in a detailed manner and tables are included to clarify the outcome. Detailed safety data can be, for instance, descriptions of serious adverse events, pregnancies, deaths or drug discontinuations due to an adverse event. In addition, safety

is divided into different sub-topics, for example, lab value measurements, extent of exposure and safety in special patient groups. These are all data that allow for extensive reporting.

From a writing point of view, there might be several reasons why the module 2.5 and 2.7 documents exceeded the recommended page length. For module 2.5 documents it might be due to the relatively new version of the CTD guidance (M4E R2). The new version demands an increase of the benefit-risk content, i.e., more repetition and expansion of the section is required, which causes the original 30-page recommendation to be rather outdated. A reason why the module 2.7 documents exceeded the page number recommendations could be because it is rather difficult to create a brief document that must contain enough information to show efficacy and safety of a drug.

A follow up question would be whether the guidance should be updated as only a few sponsors are following the 30-page and 50- to 400-page recommendation or should all sponsors reconsider how to present data in the module 2.5 and 2.7 documents to reduce and reach the preferred number of pages?

The number of tables followed the same trend as the number of pages for each module. The number of tables was lowest for the module 2.5 documents and increased continuously from module 2.7.1 to 2.7.4. The number of pages did not necessarily increase because of the high number of tables, however, the results showed that several of the documents with the greatest number of pages also contained the greatest number of tables. In other words, it is possible that fewer tables could result in a somewhat shorter document.

According to the thesis results, less table repetition did not necessarily indicate a shorter document and the reason for this might be that the representation of category 2 and 3 were similar to each other, i.e., while the inclusion criteria of tables in category 1 and 4 were clear, there was a fine line between inclusion of tables in category 2 and 3. If there were additional categories with detailed descriptions the results would show clearer differences between the categories and the number of pages. Unfortunately, as the results were collected by the author alone, time did not allow more detailed categorization of the tables. The results might depend on other factors as well. If repetition between different sections

occurs, the number of pages might stay higher despite minimum repetition of table content in the text.

The figures did not follow the same trend as the tables. The results showed that module 2.7.2 had the highest number of figures compared to the other modules (2.7.1, 2.7.3, 2.7.4). The clinical pharmacology and modeling data that is found in module 2.7.2 documents are usually better presented in figures, which will help the reader to understand the data. On the other hand, a reason why figures are not presented in such a high number in other modules might be because it is easier to create summary tables with numbers than to create clear and reader-friendly figures of the given data. An exception here might be laboratory value data, which are more easily presented as figures. However, documents that had a higher number of figures did not necessarily have a smaller number of tables. In other words, tables were not always exchanged for figures in module 2.7.2, but both figures and tables could be expressed in a high volume for one specific document. The combination of both tables and figures increased the total number of pages in several documents. A general way to reduce the document length would be to reassess how data is presented and to include fewer figures and tables. Prior reducing the number of tables or figures, a detailed discussion with the authority is required to understand what the authority is expecting from the sponsor and to ensure that essential details are included in the application dossier. This discussion is organized via a pre-submission briefing book and in a pre-submission meeting.

The CTD standards were generally followed for the level-1 headings and subheadings. However, a number of variations were identified between the documents of different sponsors and some documents had included additional level-1 headings. Additional subheadings were included for all documents, therefore increasing their length with supplementary details. When writing the documents, it is always important that they are structured in a way that it is easy to guide the reader through all the sections. As stated in the CTD guideline "*Throughout the Common Technical Document, the display of information should be unambiguous and transparent, in order to facilitate the review of the basic data and to help a reviewer become quickly oriented to the application contents"* (EMA, 2004). It is implied that having a long document with an unclear structure might

be difficult for the reader to follow. Subheadings are included to ease the reader through a complex and detailed document and allow the reviewer to concentrate on one specific topic at a time. As the results showed, the high numbers of subheadings are vastly used by all of the sponsors, compared to the number of subheadings recommended by the CTD. Sponsors might have the impression that a vast number of subheadings are necessary as the number of subheadings recommended by the CTD might lead to poorly structured documents. This may be a reason why additional subheadings are included. On the other hand, additional subheadings might in turn lead to inclusion of non-essential details or repetition of data. Writing documents with a length closer to the CTD recommendations might be possible if the level-1 headings and subheadings were divided according to the CTD guideline and if only information that is specified for these specific documents was included.

There might be many reasons why additional content (in addition to the CTD recommendations) is included in the documents. When writing strategies are discussed it is essential to consider what data is vital to present in the documents. The writing approaches are thoroughly discussed with the health authority, and essential data is identified. In this occasion, the CTD recommendations are secondary as the meetings with the authority will set the base for the key points. Implementation of the essential information might lead to use of more level-1 headings and subheading with the intention to create a reader-friendly document. Additionally, there are other recommendations to be followed except the CTD guideline, e.g., recommendations based on the number of summarized studies and how to structure them. There may also be a reason to mention the CTD recommendation itself that was introduced long time ago and might be rather outdated.

There are also other factors that will influence both the document length and how well the CTD recommendations are followed. The planning of the trial itself, the chemical entity and the product type (e.g., generic or orphan) are example of these factors. For instance, today there are more rules to be followed when performing the clinical trials. Chemical entities are also more complex than before, and more difficult descriptions and detailed reporting are required. In addition, it is more challenging to develop new drugs based on

small molecules and development of biological drugs have become widespread in the pharmaceutical industry, which requires even further detailed clarifications to ensure safety and efficacy.

6.2 Sponsor differences

Individual writing techniques could be a result of variation within a specific sponsor as there are several medical writers interpreting the results. Particular medical writers might favor the inclusion of supplementary information to make it clear for the reader, while other medical writers may try to include as little information as possible. It is also important to remember that what is written in the documents does not only depend on the medical writers but the whole study team. There might be sponsor-specific, country-specific, or even individual preferences to explain what information should be included, i.e., senior experts might have created their own perspective of what to include in the documents, which could lead to differences in each document.

While a fast marketing authorization is the key goal for each sponsor, the writing strategies to reach the goal may differ. For some sponsors the strategy is built up to include extensive amount of content and repetition to ensure that the endpoints and benefits are clearly understood. Other sponsors may concentrate on a clear message with short descriptions and less repetition.

6.3 What have we learned?

If guidelines with greater details were implemented, a clearer understanding of how data should be presented would arise and there would be less room for misinterpretation. Simplifying the documentation process could impact resources, quality and, in addition, it would be crystal clear how the authorities should interpret data. When documents are simplified, less repetition is found in the text. When less repetition is found in the text, there is a reduced risk that errors occur when updates in the text are required. Besides repetition, it is also important to understand how to present data. Individual patient data always have a higher chance of being reidentified, resulting in an increased need for redaction resources, whereas redaction resources are reduced if data are summarized.

Not only the sponsor's resources but also patients would benefit from document simplification. There would be increased trust between the public and the sponsor, as unnecessary information that might lead to reidentification of a particular person who participated in the trial would be excluded. It is not well defined which simplification approaches apply to which sponsor, but it is obvious that all sponsors should examine their approaches and consider whether the documents could be simplified. Nevertheless, the collected data clearly show that there is room for improvement for all sponsors regarding the document length. Improvements could be, e.g.:

- To simplify the documentation at the start of the writing process
- To follow the CTD recommendations carefully and narrow down the content accordingly
- To implement fewer explanatory figures and tables, especially tables
- To reduce repetitive writing processes and create a better relationship between tables and text
- To implement number of level-1 headings and subheadings in accordance with the CTD
- To delete out-of-scope text for a specific document
- To place information, if possible, in sections that will not be publicly available
- To reduce repetition by mentioning the same conclusions, paragraphs or table content once or as few times as possible, which would shorten the total document length
- To include cross-references where possible instead of repeating what is said elsewhere. Highest potential for cross-referencing is into module 5 (Clinical Study Reports) documents

Improvements would not only be beneficial for the sponsor but also the regulatory authorities, e.g.:

- To update the CTD guideline by taking into consideration the new transparency guidelines
- To make more detailed recommendations in the guidelines to create less sponsor variation

- ❖ Regulatory authorities should be more transparent with what kind of information must be included in the MAA

6.4 Other thoughts

It should be considered whether the sponsors should modify the writing strategies as the regulations and recommendations are changing, in order to keep up with transparency measures. Another concern is whether it is possible to keep up with the changes and whether the recommendations (e.g., the CTD guideline) take into consideration that some parts of the modules are to be made publicly available.

EMA has certain recommendations for how PPD (Protected Personal Data) and CCI (Commercially Confidential Information) should be handled, i.e., redaction and anonymization recommendations. According to EMA, PPD is preferred to be anonymized where possible, however, it is the sponsor's responsibility to clarify the PPD and the calculated risks of re-identification must be provided in the anonymization report. EMA also provide some guidance for direct and indirect identification of an individual who participated in a trial. In accordance with EMA, CCI often must be redacted and a justification table is used as a communication tool where the sponsor clarifies the reasonings for CCI redactions in each document. The justifications are either accepted, partially accepted or rejected by EMA (or HC). Further explanations can be provided by the sponsor if needed but the final conclusion is formed by EMA. The sponsor can appeal to the Court of Justice of the European Union in case an agreement concerning CCI cannot be reached.

Since the authorities ultimately decide which redaction and anonymization suggestions are approved - are the authorities sympathetic to sponsors considering the burden of potentially releasing patient and company information or do they focus purely on what they believe to be important?

A whole new discussion could also be focused on the recommendations. There are numerous initiatives that try to include additional countries for a harmonized document structure, which should lead to similar structures and easily readable documents. Stronger harmonization than what we have today would reduce interpretation, and details like how

to repeat, what to repeat and how many figures and tables to include could become clearer. The ideal solution would be to have stricter rules that would cover the grey area. Resources would be reduced, as it would be easier to follow the guidance and less discussion of how to present data would be required. The question remains whether recommendations can create strong harmonization between sponsors or whether concrete directions should be put in place.

7 Further investigations

In order to draw conclusions between document structures and specific indications, it would require several different analyses on an increased scale with specific indications and document types. However, today there are not enough module 2 packages to be compared and a more thorough analysis would have to wait a couple of years until additional submission packages have been published on the respective websites of EMA and HC.

8 Summary in Swedish - Svensk sammanfattning

Potential för förenkling av skrivandet av inlämningsdokument för läkemedel: Utvärdering av allmänt tillgängliga modul 2-dokument från olika läkemedelsföretag

Detta projekt handlar om möjligheterna att förenkla skrivandet av kliniska modul 2.5- och 2.7-dokument som är en del av de dokument som blir offentligt tillgängliga efter ansökan om ett marknadstillstånd. Tillgängligheten är ett resultat av de transparenta regler vi har idag.

1 Introduktion

Innan läkemedlet kan börja användas måste det gå igenom olika faser av säkerhets- och effektivitetssprövningar. Dessa delas in i pre-kliniska och kliniska prövningar.

I pre-kliniska prövningar fastställs de första toxikologiska och farmakologiska testerna samt säkerhetsprövningar. Det är en förberedelse för att hitta ett säkert läkemedel samt en säker dos för människan i de efterföljande kliniska prövningarna. Data från dessa pre-kliniska prövningar skickas in till myndigheterna som fattar beslut kring om prövningarna kan gå vidare till följande fas, de kliniska prövningarna. (David J. Kerr, 2006; E.L. Andrade, 2016).

Kliniska prövningar utförs på människor (både friska och sådana med ifrågavarande sjukdom) för att undersöka det bästa sättet att diagnostisera och behandla en sjukdom. Prövningarna fokuserar på läkemedlets effektivitet, säkerhet och toxicitet samt att hitta de rätta och mest effektiva läkemedelsdoserna. Data från dessa prövningar skickas återigen in till myndigheterna som beslutar om ett läkemedelstillstånd kan utfärdas. (David J. Kerr, 2006).

2 CTD - Common Technical Document

Alla dokument som skickas in till läkemedelsmyndigheten måste följa vissa strukturer. Ett regelverk som många länder följer är ett dokument som heter "Common Technical Document" (CTD). CTD-riktlinjer är rekommendationer som upprätthålls av ICH

(International Conference on Harmonisation) och består av fem olika delar: modul 1, modul 2, modul 3, modul 4 och modul 5. Varje modul innehåller detaljerade instruktioner om hur dokumenten bör formateras (t.ex. specifika rubriker, textstil, marginaler) och i vilka sektioner specifik information bör finnas. (EMA, 2003; EMA, 2004; EMA, 2016).

Alla läkemedelsföretag som följer CTD-standarderna har samma dokumentstruktur, vilket i sin tur leder till att skrivprocessen blir mindre tidskrävande och såvida minskar resursbehovet. En global standard gör det lättare att läsa och hitta information, både för tillsynsmyndigheterna och läkemedelsindustrin. (EMA, 2004).

Modul 1 innehåller administrativ information och förskrivningsinformation som i sig själv inte räknas in i CTD. Modul 2 är slutsatser och sammanfattningar av modul 3, 4 och 5. Modul 3 innehåller information om kvalitet, modul 4 handlar om preklinisk säkerhet (så kallade toxikologiska prövningar) och modul 5 handlar om klinisk effektivitet. (EMA, 2003).

Alla moduler är vidare uppdelade i olika avsnitt och delområden. Alla dessa moduler är alltså del av marknadsföringstillståndet för en läkemedelsprodukt och av dessa måste (som en följd av de nyaste transparensreglerna) modul 2-och modul 5-dokument publiceras för allmänheten. Eftersom denna avhandling handlar om allmänt tillgängliga modul 2-dokument kommer även fokus att ligga på dessa. Modul 2-dokument uppdelas i sju olika delar:

- (2.1) CTD - Innehållsförteckning
- (2.2) CTD - Introduktion
- (2.3) Övergripande kvalitetssammanfattning
- (2.4) Icke-klinisk översikt
- (2.5) Klinisk översikt
- (2.6) Icke-kliniska skriftliga och tabulerade sammanfattningar
- (2.7) Klinisk sammanfattning. (EMA, 2003).

Av dessa delar är de allmänt tillgängliga dokumenten endast den kliniska översiktsmodulen-2.5 och den kliniska sammanfattningsmodulen-2.7. Dessa delas vidare upp i sina egna avsnitt och delområden:

- Modul 2.5 - Klinisk översikt
- Modul 2.7.1 - Klinisk sammanfattning av biofarmaceutiska prövningar och analysmetoder
- Modul 2.7.2 - Klinisk sammanfattning av kliniska farmakologiska prövningar
- Modul 2.7.3 - Klinisk sammanfattning av klinisk effekt
- Modul 2.7.4 - Klinisk sammanfattning av klinisk säkerhet. (EMA, 2016).

Dessa moduler har sedan sina egna huvud-och underrubriker. Huvudrubrikerna hittas i kapitel 2.1, Table 2.

3 Ansökan om läkemedelstillstånd

Det flesta nya läkemedel i Europa godkänns genom ett så kallat centraliserat förfarande. Detta betyder att ett läkemedelstillstånd kan sökas en gång och därmed godkännas i alla europeiska medlemsstater tillika. Fördelen med detta förfarande är en centraliserad säkerhetsövervakning samt produktinformation på alla EU-språk. När ett läkemedels-tillstånd ansöks om hos den europeiska läkemedelsmyndigheten (EMA) påbörjas en högst 314 dagar lång granskningsprocess som slutar i ett godkännande eller avkastande av läkemedlet. En del av de dokument som hör till läkemedelstillståndsansökan kommer senare bli publicerade för allmänheten och detta är ett resultat av transparensen som krävs inom läkemedelsindustrin. (EMA, 2019)

4 Transparens i Europa och Kanada

När det gäller kliniska läkemedelsprövningar är de främsta orsakerna för transparens:

- Att skapa förtroende mellan allmänheten och läkemedelsindustrin
- Att ge möjligheten för vem som helst att analysera olika prövningar och att skapa en personlig uppfattning om en viss produkt
- Att tvinga fram innovativt tänkande och skapa forskningsmöjligheter för andra företag (European Clinical Research Infrastructure (ECRIN), n.d.; EMA, n.d. (d))

Ett flertal riktlinjer gällande transparens har skrivits och tillämpats under årens lopp. Kontinuerliga uppdateringar och nya riktlinjer har lett till var vi står idag. Kraven och riktlinjerna kommer kontinuerligt att utvecklas och det slutliga resultatet av denna implementering förväntas leda till publicering av alla kliniska prövningar, oavsett resultat.

(European Clinical Research Infrastructure (ECRIN), n.d.; EMA, n.d. (d); Health Canada, Public Release of clinical Information: guidance document, 2019a).

- **EMA Policy 0070 och Health Canada PRCI**

De senaste riktlinjerna för transparens har implementerats i Europa (EMA Policy 0070) och Kanada (Health Canada PRCI) under åren 2015 och 2019. Dessa riktlinjer har gjort det möjligt för detta projekt att existera, eftersom modul 2.5- och 2.7 dokument, som är en del av läkemedelstillståndet, nu är allmänt tillgängliga. (European Clinical Research Infrastructure (ECRIN), n.d.; EMA, n.d. (d); Health Canada, Public Release of clinical Information: guidance document, 2019a).

Riktlinjerna innehåller information om vilka dokument som bör publiceras offentligt i Europa, respektive Kanada. Dessa riktlinjer ger även vägledning i hur anonymisering och redigering av företagshemligheter och persondata bör utföras i de olika dokumenten. (European commission, 2015; Health Canada, 2019a).

5 Målsättning

Målet med denna avhandling är att analysera möjliga förslag för att förenkla utarbetandet av offentligt tillgängliga modul 2-dokument genom att:

- Jämföra strukturella skillnader i 20 olika modul 2-dokument för 10 olika sponsorer[13]
- Jämföra mängden datarepetition mellan texttabeller och textstycken i modul 2-dokument

Syftet är att klargöra om omfattande variationer existerar mellan sponsorernas dokument och skrivstrategier. De nuvarande skrivprocesserna är tids- och resurskrävande och det skulle vara fördelaktigt att kunna eliminera överflödig och upprepad information i dokumenten.

Avsikten är att göra alla läkemedelsföretag medvetna om resultaten och hur rekommendationerna kan implementeras i det dagliga arbetet.

[13] En sponsor är det ledande företaget som ekonomiskt stödjer en aktivitet eller ett evenemang, ofta ett läkemedelsföretag

6 Material och metoder

- **Litteraturöversikt**

En litteraturöversikt av CTD, läkemedelstillståndsprocessen och transparensriktlinjer är nödvändig för att skapa en förståelse av dessa modul 2-dokument.

CTD beskriver reglerna för dokumenteringsstrukturen som krävs för ett marknadstillstånd. Läkemedelstillståndsprocessen klargörs genom beskrivning av prekliniska prövningar till ett marknadstillstånd och transparentriktlinjerna anger tydligt vad som menas med transparens, dess nödvändighet och hur detta ska tillämpas i de olika dokumenten.

- **Sponsorer och dokument**

Inom ramen för denna forskning har författaren valt att undersöka 10 sponsorer. För varje sponsor har valts två allmänt tillgängligt modul 2-dokumentpaket. Ett paket från EMA:s webbsida och ett paket från Health Canadas webbsida. Tillsammans blir det en full analys av 10 olika sponsorer och 20 modul 2-dokument. Ett modul 2-dokumentpaket avser de allmänt tillgängliga dokumenten i modulavsnitten: 2.5, 2.7.1, 2.7.2, 2.7.3 och 2.7.4.

- **Variabler**

Vid tidpunkten för dokumentuppladdning registrerades följande information:

- Indikation, ACT kod, först-i-klass läkemedel, publiceringsår och dokumentationstyp

För en kvantitativ analys jämfördes följande variabler:

- Antalet sidor per dokument
- Antalet tabeller
- Antalet figurer
- Antalet huvudrubriker och underrubriker
- Indelning av tabeller i fyra olika kategorier, beroende på hur mycket av informationen i tabellerna som repeteras i textformat (endast analys av modul 2.5- och 2.7.4-dokument)

7 Resultat

CTD-riktlinjer är endast rekommendationer och strukturvariationer från dessa rekommendationer betyder inte att dokumenten är inkorrekta.

Enligt CTD-strukturen bör modul 2.5-dokumenten vara omkring 30 sidor långa. Av alla modul 2.5-dokument hade 95 % åtminstone 50 sidor långa dokument och därav följdes inte CTD-rekommendationen. Rekommenderad sidolängd för det kliniska sammanfattningsdokumentet 2.7 (modul 2.7.1, 2.7.2, 2.7.3 och 2.7.4) bör vara runt 50 till 400 sidor. Endast 55 % av dokumenten uppfyllde denna rekommendation.

Enligt CTD-rekommendationerna föredras tabeller och figurer i dokumenten när det kan förbättra läsbarheten men det är dock upp till sponsorn att bestämma när information och resultat ska presenteras i sådan form för att ge bättre klarhet. Detta är en vag rekommendation som resulterar i olika tolkningar, vilket även framgick i datainsamlingen som tydligt visade att mängden figurer och tabeller varierade mellan olika sponsorer.

Alla sponsorer följde CTD-strukturen gällande huvudrubrikerna men mängden information och titelnamn som presenterades i varje avsnitt varierade både mellan dokumentmodulerna och sponsorerna. En del sponsorer inkluderade dock fler huvudrubriker än vad som rekommenderas av CTD. Det är tydligt att alla sponsorer tolkade riktlinjerna på sitt eget sätt och en gråzon med utrymme för egna slutsatser var synlig. Detta var ett resultat som visade att resursanvändningen varierarar mellan sponsorerna eftersom vissa sponsorer inkluderar mer data i dokumenten än andra.

Enligt CTD bör modul 2.5 inkludera 7 underrubriker vilket ingen sponsor följde eftersom alla hade inkluderat 21 eller fler underrubriker. CTD har inga specifika krav på underrubriker för modul 2.7.1 och 2.7.2 men majoriteten av alla dokument överskred 10 eller fler underrubriker. Module 2.7.3 inkluderade 3 underrubriker och modul 2.7.4 inkluderade 15 underrubriker. Endast ett dokument följde rekommendationen för modul 2.7.3 och inga dokument följde rekommendationen för modul 2.7.4. Båda modulerna (2.7.3 och 2.7.4) överskred underrubriksmängden med upp till 105 respektive 138 underrubriker.

Det finns inte några specifika rekommendationer om hur man bör upprepar data från tabeller i textformat men det är underförstått att upprepning av något slag bör undvikas. Repetition av data från tabeller i textformat skedde i olika grad men en modulspecifik trend var synlig. I modul 2.5-dokumenten ansågs det mindre viktigt att upprepa data från tabeller i textformat medan mer upprepning föredrogs i modul 2.7.4-dokumenten. Eftersom modul 2.5-dokumenten endast är en översikt och modul 2.7.4-dokumenten en sammanfattning var också dessa resultat förväntade. Även om det var lätt att upptäcka en trend mellan de olika modulerna, fanns det fortfarande en individuell skillnad mellan sponsorerna.

8 Slutsatser

Mer detaljerade riktlinjer skulle resultera i en tydligare förståelse för hur data ska presenteras och gråzonen för tolkningsutrymme skulle försvinna. En förenklad dokumentations-process kan påverka resurserna och det skulle bli tydligare för myndigheterna hur data bör tolkas. Mindre upprepning betyder mindre risk för misstag i de fall då dokumenten bör uppdateras. Förutom upprepning är det också viktigt att förstå vilket det bästa sättet är att presentera data. Enskilda patientdata har alltid en högre chans till patientidentifiering, vilket resulterar i ett större behov av redigerings- och anonymiseringsresurser. Om data däremot presenteras på en sammanfattande nivå minimeras dessa resurser.

Dokumentförenkling skulle inte bara vara till nytta för sponsorerna men ur ett patientperspektiv skulle ett ökat förtroende mellan allmänheten och sponsorerna skapas, i och med att fokus skulle ligga på att exkludera onödig information som kan leda till identifiering av en specifik patient.

Rekommenderade förbättringar kunde till exempel vara:

- Att förenkla dokumentationen i början av skrivprocessen
- Att implementera färre figurer och tabeller
- Att minska på repetitiva skrivprocesser och skapa en bättre relation mellan text och tabeller

- Att implementera färre underrubriker och mindre text, för att på detta sätt förkorta dokumentlängden
- Att radera text som är utanför omfånget för ett visst dokument
- Att placera information (i den mån det är möjlighet) i moduler som inte är allmänt tillgängliga
- Att nämna slutsatser, paragrafer eller tabellinnehåll endast en gång, vilket skulle leda till mindre upprepning samt förkortning av den totala dokumentlängden
- Att använd korsreferenser där det är möjligt istället för att upprepa vad som redan nämnts i en annan modul eller i ett annat avsnitt i samma dokument

9 References

Bayer, Personal Communication. (2021, Mar 29).

Canada, G. o. (2021). *Drug and health products legislation and guidelines*. Retrieved Feb 22, 2021, from Overview of Vanessa's Law: https://www.canada.ca/en/health-canada/services/drugs-health-products/legislation-guidelines/overview-vannessa-law-protecting-canadians-unsafe-drugs-act-vanessa-law-amendments-food-drugs-act.html

David J. Kerr, K. K. (2006). Clinical Trials Explained. In *Clinical trials explained : a guide to clinical trials in the NHS for healthcare professionals* (pp. 4-17). Blackwell.

E.L. Andrade, A. B. (2016). Non-clinical studies required for new drug development - Part 1: early in silico and in vitro studies, new target discovery and validation, proof of principles and robustness of animal studies. *Brazilian Journal of Medical and Biological Research, 49(11)*, 1-9. doi:http://dx.doi.org/10.1590/1414-431X20165644

EDSP (European Data Protection Supervisor). (1995, Nov 23). *Directive 95/46/EC.* Retrieved Jan 25, 2021, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:31995L0046

EMA. (2003, Jul). *ICH Topic M 4 E, Common Technical Document for the Registration of Pharmaceuticals for Human - Efficacy. CPMP/ICH/2887/99*. Retrieved Jan 5, 2021, from https://www.ema.europa.eu/en/documents/scientific-guideline/ich-m-4-e-common-technical-document-registration-pharmaceuticals-human-use-efficacy-step-5_en.pdf

EMA. (2003). ICH Topic M4Q Common Technical Document for the Registration of Pharmaceuticals for Human Use, Quality Overall summary of module 2 and module 3: Quality (CPMP/ICH/2887/99). EMA. Retrieved Jul 24, 2020, from https://www.ema.europa.eu/en/documents/scientific-guideline/ich-m-4-q-

common-technical-document-registration-pharmaceuticals-human-use-quality-step-5_en.pdf

EMA. (2004). ICH Topic M4, Common Technical Document for the Registration of Pharmaceuticals for Human Use – Organisation Common Technical Document. (CPMP/ICH/2887/99). Retrieved Jul 24, 2020, from https://www.ema.europa.eu/en/documents/scientific-guideline/ich-m-4-common-technical-document-registration-pharmaceuticals-human-use-organisation-ctd-step-5_en.pdf

EMA. (2007, Feb). VOLUME 2A Procedures for marketing authorisation CHAPTER 2 Mutual Recognition - ENTR/F2/SM (2007). Retrieved Mar 26, 2021, from https://ec.europa.eu/health//sites/health/files/files/eudralex/vol-2/a/vol2a_chap2_2007-02_en.pdf

EMA. (2015). The Centralised Procedure at the EMA. Retrieved Feb 23, 2021, from https://www.ema.europa.eu/en/documents/presentation/presentation-centralised-procedure-european-medicines-agency_en.pdf

EMA. (2016). M4E(R2) - Common technical document for the registration of pharmaceuticals for human use – Efficacy. Retrieved Aug 11, 2020, from https://www.ema.europa.eu/en/documents/scientific-guideline/ich-m4e-r2-common-technical-document-registration-pharmaceuticals-human-use-efficacy-step-5_en.pdf

EMA. (2018). External Guidance on the Implementation of the European Medicine Agency Policy on the Publication of Clinical Data for Medical Products for Human Use (EMA/90915/2016). (4). Retrieved Aug 13, 2020, from https://www.ema.europa.eu/en/human-regulatory/marketing-authorisation/clinical-data-publication/support-industry/external-guidance-implementation-european-medicines-agency-policy-publication-clinical-data

EMA. (2019, Mar 4). Authorisation of medicines. Retrieved Feb 2, 2021, from https://www.ema.europa.eu/en/about-us/what-we-do/authorisation-medicines

EMA. (2019). European Medicines Agency policy on publication of clinical data for medicinal products for human use. (EMA/144064/2019). Retrieved Jul 24, 2020, from https://www.ema.europa.eu/en/documents/other/european-medicines-agency-policy-publication-clinical-data-medicinal-products-human-use_en.pdf

EMA. (2021a). *About the EU Clinical Trials Register*. Retrieved Feb 22, 2021, from EU Clinical Trials Register: https://www.clinicaltrialsregister.eu/about.html

EMA. (2021b). Clinical data publication. Retrieved Mar 1, 2021, from https://www.ema.europa.eu/en/human-regulatory/marketing-authorisation/clinical-data-publication

EMA. (n.d. (a)). Committees, working parties and other groups. Retrieved May 5, 2021, from https://www.ema.europa.eu/en/committees-working-parties-other-groups

EMA. (n.d. (b)). Medicines for use outside the European Union. Retrieved Apr 28, 2021, from https://www.ema.europa.eu/en/human-regulatory/marketing-authorisation/medicines-use-outside-european-union

EMA. (n.d. (c)). Obtaining an EU marketing authorisation, step-by-step. Retrieved Aug 12, 2020, from https://www.ema.europa.eu/en/human-regulatory/marketing-authorisation/obtaining-eu-marketing-authorisation-step-step

EMA. (n.d. (d)). Transparency. Retrieved Aug 12, 2020, from https://www.ema.europa.eu/en/about-us/how-we-work/transparency

EudraCT. (n.d.). *Welcome to the EudraCT public home page*. (E. Commission, Editor) Retrieved Feb 22, 2021, from EudraCT: https://eudract.ema.europa.eu/

European Clinical Research Infrastructure (ECRIN). (n.d.). Transparency in Clinical Research. Retrieved Aug 12, 2020, from https://ecrin.org/node/583

European commission. (2015). Volume 2A, Procedures for marketing authorisation, chapter 1 - Marketing authorisation. 9. Retrieved May 5, 2021, from https://ec.europa.eu/health/sites/default/files/files/eudralex/vol-2/vol2a_chap1_rev6_201612.pdf

FDA. (2018). *Milestones in U.S. Food and Drug Law History*. Retrieved Feb 18, 2021, from U.S. Food & Drug Administration: https://www.fda.gov/about-fda/fdas-evolving-regulatory-powers/milestones-us-food-and-drug-law-history

FDA. (2019). New Drug Application (NDA). Retrieved Aug 12, 2020, from https://www.fda.gov/drugs/types-applications/new-drug-application-nda

FDA. (2020, Mar 26). *U.S Food and Drug Administration - Clinical Data Summary Pilot Program.* Retrieved Sep 7, 2020, from https://www.fda.gov/drugs/development-approval-process-drugs/clinical-data-summary-pilot-program

Health Canada. (2019a). Public Release of clinical Information: guidance document. Retrieved Jul 24, 2020, from https://www.canada.ca/en/health-canada/services/drug-health-product-review-approval/profile-public-release-clinical-information-guidance/document.html

Health Canada. (2019b). Guidance Document: The Management of Drug Submission and Applications. Retrieved Aug 12, 2020, from https://www.canada.ca/en/health-canada/services/drugs-health-products/drug-products/applications-submissions/guidance-documents/management-drug-submissions/industry.html#a5.2

ICH. (1995). ICH Harmonised Tripartite guideline. Structure and Contenct of Clinical Study Reports E3. (4). Retrieved Aug 13, 2020, from https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-3-structure-content-clinical-study-reports-step-5_en.pdf

ICH. (n.d.). *ICH guidelines*. Retrieved Oct 7, 2020, from https://www.ich.org/page/ich-guidelines

IFPMA, PhRMA, efpia, & JPMA. (2018, January 15). *Joint Position on the Disclosure of Clinical Trial.* Retrieved Sep 7, 2020, from https://www.ifpma.org/wp-content/uploads/2010/11/Joint-Position-on-Disclosure-of-CT-Info-via-CT-Registries-Revised-Jan2018-vFINAL.pdf

Information and Privacy Commissioner of Ontario. (2016). *De-identification Guidelines for Structured Data.* Retrieved Dec 1, 2020, from https://www.ipc.on.ca/wp-content/uploads/2016/08/Deidentification-Guidelines-for-Structured-Data.pdf

Japan Pharmaceutical Manufacturers Association. (2018). *Pharmaceutical Administration and Regulations in Japan. Information on Japanese Regulatory Affairs. Regulatory Information Task Force Japan Pharmaceutical Manufacturers Association.* Retrieved Sep 7, 2020, from http://www.jpma.or.jp/english/parj/pdf/2018.pdf

Jordan, D. (2014). An overview of the Common Technical document (CTD) regulatory dossier. *The European Medical Writers Association, 23*(2), 101-105. doi:DOI: 10.1179/2047480614Z.000000000207

NIH. (n.d.). ClinicalTrials.gov is a database of privately and publicly funded clinical studies conducted around the world. (U. N. Medicine, Ed.) Retrieved May 5, 2021, from https://clinicaltrials.gov/ct2/home

Office of the Privacy Commissioner of Canada. (2019). The Privacy Act in Brief. Retrieved Aug 26, 2020, from https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-privacy-act/pa_brief/#s02

Patrick Waller, M. H.-W. (2017). *An introduction to pharmacovigilance* (2nd ed.).

PMDA. (n.d.). *Pharmaceuticals and Medical Devices Agency - Outline of Reviews and Related Services.* Retrieved Sep 7, 2020, from https://www.pmda.go.jp/english/review-services/outline/0001.html

The European Union. (2004). Regulation (EC) No 726/2004 of the European Parliament and of the Council, article 58. *EUR-Lex*. Retrieved Dec 5, 2020, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32004R0726

The European Union. (2008, Nov 24). Commission regulation (EC) No 1234/2008. Retrieved Sep 5, 2020, from https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:334:0007:0024:en:PDF

The European Union. (2016a, Apr 27). *Official Journal of the European Union. REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. Chapter I. Definitions, article 4.* Retrieved Nov 18, 2020, from https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN#d1e1489-1-1

The European Union. (2016b, Apr 27). *Official Journal of the European Union. REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. Chapter II Principles, article 5.* Retrieved Nov 18, 2020, from https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN#d1e1807-1-1

WHO. (2018, Jul 19). *Bulletin of the World Health Organization*. Retrieved Mar 25, 2021, from Increasing transparency and accountability in national pharmaceutical systems: https://www.who.int/bulletin/volumes/96/11/17-206516/en/

Vuolo, M. (2020, Dec 2). Breaking the Document Mindset. *Pharmaceutical Engineering Magazine - ISPE*. Retrieved Apr 8, 2021, from https://ispe.org/pharmaceutical-engineering/ispeak/breaking-document-mindset