PhUSE US Connect 2019
Paper SA01
Paper SA01

EMA, Health Canada, FDA and PMDA: Four agencies tackle Data Sharing. Synergies and Differences

Jean-Marc Ferran, Qualiance ApS, Copenhagen, Denmark
Liz Roberts, UCB, Washington DC, USA

ABSTRACT

The European Medicines Agency (EMA) and then Health Canada have respectively launched separate data transparency policies related to the public posting of submission documents. The two regulators are providing additional guidance to help with implementation of their policies and each formed stakeholders' groups during 2017.

These two stakeholders' groups, EMA Technical Anonymisation Group (TAG) and Health Canada's Reference Group (HCRG), include a variety of professionals and representatives across Life Sciences and focus on topics around data anonymization, risk of re-identification and data utility while HCRG also addresses scope and process- related matters.

This presentation will elaborate on the approaches, goals and achievements of these two stakeholders' groups and reflect on the EMA and Health Canada data sharing processes and requirements and where they overlap and differ from each other. In addition, the recent initiative and pilots from FDA as well as related activities within PMDA, Japan will also be discussed.

INTRODUCTION

This paper intends to provide an exhaustive summary of clinical data transparency efforts from EMA, Health Canada, FDA and Japan, how each agency engaged with stakeholders and how their initiatives differ.

The Clinical Data Transparency landscape has greatly changed since the first talks initiated at the EMA in 2012 and today, where companies provide qualified researchers with access to anonymized Individual Patient Data (IPD) upon request, EMA Policy 0070 has been active since the beginning of 2015, and other agencies are discussing similar initiatives. The current regulatory landscape focuses on the publication of submission documents including those within Module 2 and Module 5. Regulatory discussions around the sharing of IPD are postponed to a later time with no firm commitments on timelines.

While sponsors are working hard to adapt their processes, form teams and keep up to date with regulations and policies, it appears that the different agencies are taking different directions when it comes to requirements that apply to same documents. Sponsors are trying to understand both similarities and differences to be able to provide de-identified documents that are compliant with the different requirements while using their resources wisely.

This paper explores similarities and differences in terminology, processes and de-identification requirements that could be useful to highlight to both help regulators align, where feasible, to alert sponsors to competing requirements, especially when considering the risk of re-identification where the same documents would be globally available in different de-identified formats in different regions.

AGENCIES ENGAGING WITH STAKEHOLDERS

All four agencies are engaging in different ways and to different extents with stakeholders, including academia, patient organizations, pharmaceutical companies, data anonymization subject matter experts (SMEs), software vendors, and standards organizations, etc. through activities such as:

- Public consultations
- Stakeholders consultation
- Dedicated working groups or
- Pilots with industry

EMA PUBLIC AND STAKEHOLDERS GROUPS CONSULTATION & EXTERNAL GUIDANCE

EMA started their initiative to develop a clinical data publication policy with a workshop on 22 November 2012. Following the event, the Agency issued a call for nominations to join advisory groups to inform the policy on five topics and form five Clinical Trial Advisory Groups (CTAG): Protecting Patient Confidentiality, Clinical Trial Data Formats, Rules of Engagement, Good Analysis Practice, and Legal Aspects. Each of the five CTAGs issued an “Advice” document [0].

The CTAGs included 318 individuals with 97 from Industry, 65 from Academia, 23 representing Healthcare professionals and 19 representing Patients’ organizations.

In June 2013, EMA released a draft policy on publication and access to clinical data for a 3-month public consultation [1]. The initial draft policy was covering both what we know as Phase 1 (submission documents) and Phase 2 (IPD), and EMA received over 1,000 comments. Following this, the EMA held meetings (internal, and with stakeholders) in 2014, which led to the release of *European Medicines Agency policy on publication of clinical data for medicinal products for human use* of 02 October 2014, better known as *Policy 0070*, with effective date of 01 January 2015 focusing on disclosure of submissions documents (Phase 1 of the policy). A stepwise implementation of the policy was described with plans for IPD sharing to follow (Phase 2 of the policy) [2].

In 2015 EMA consulted stakeholders extensively, as well as the European Ombudsman and the European Data Protection Supervisor, which led to the development and finalization of the External guidance on the implementation of the EMA policy on the publication of clinical data for medicinal products for human use [3], better known as EMA Policy 0070 External Guidance. EMA provided the guidance for review to organizations that were part of their stakeholder group. Stakeholder organizations (scientific journals, pharmaceutical companies, SMEs, academics, standards organizations, and patients groups, etc.) were invited to the EMA offices or attended remotely and provided their comments directly in two separate meetings. Following these two meetings, EMA released the first version of the guidance on 02 March 2016. EMA continues to engage industry organizations from the stakeholders’ group on regular teleconferences where challenges on implementing Policy 0070 and any required clarifications are discussed. The guidance is updated on a regular basis as outlined in Table 1 where updates related to scope clarifications are highlighted.

Table 1 - EMA Policy 0070 External Guidance Updates

version	Date	Scope Clarifications
1.0	02MAR16	• New Document
1.1	21DEC16	• *Additional information concerning EMA's position on clinical reports that are submitted as part of /or cross-referred to within a regulatory application in the context of policy 0070 has been added.* • *Practical guidance for the individual patient data (IPD) out of scope sections has been included.* • *In cases where an applicant/MAH (the Transferor) transfers a MA to another company (the Transferee), the Transferee becomes responsible for the transferred product and accepts all prior agreements under policy 0070 between the Transferor and the EMA.*
1.2	12APR17	• *Clarification has been provided on which clinical reports should be submitted for publication.* • *Clarification that individual patient data listings contained in CSR section 14.3.4 "Abnormal Laboratory Value Listing" can be considered out of scope of phase 1 of Policy 0070.*
1.3	22SEP17	• *Clarification has been provided on which clinical reports should be submitted for publication with regards to extension or modification of indications in the paediatric population:* • *Inclusion of revised wording regarding the submission of duplicate marketing authorisations and inclusion of a cover letter template to be used upon the submission of the final redacted document package of duplicate applications* • *Clarification added on the definition of listings out of scope of Phase 1 of Policy 0070* • *Clarification on the instructions on how to process information considered to be out of scope of phase 1 of Policy 0070*
1.4	15OCT18	• *Clarification has been provided on the publication of withdrawn applications in cases where the application has been re- submitted or has an agreed re-submission date* • *Clarification added on the publication of clinical studies where the main period/phase of a clinical study is still on-going at the time of publication* • *In addition to the checklist for the Redaction Proposal Document package, a new checklist for the Final Redacted Document package has been added.* • *Clarification added on the labelling requirements of information considered to be out of scope of phase 1 of Policy 0070*

EMA POLICY 0070 FIRST RESULTS

After a year, the PhUSE Data Transparency Working Group analyzed the Policy 0070 submissions up to that date [4] and assessed for each submission (57 at the time), the de-identification methods used, how companies justified clinical data utility, etc.

Figure 1 displays results of anonymization methods vs de-identification techniques in general and for narratives only, and demonstrate that Policy 0070 External Guidance requirements can be challenging to be met. In particular, a majority of companies were using qualitative methods in connection with redaction.

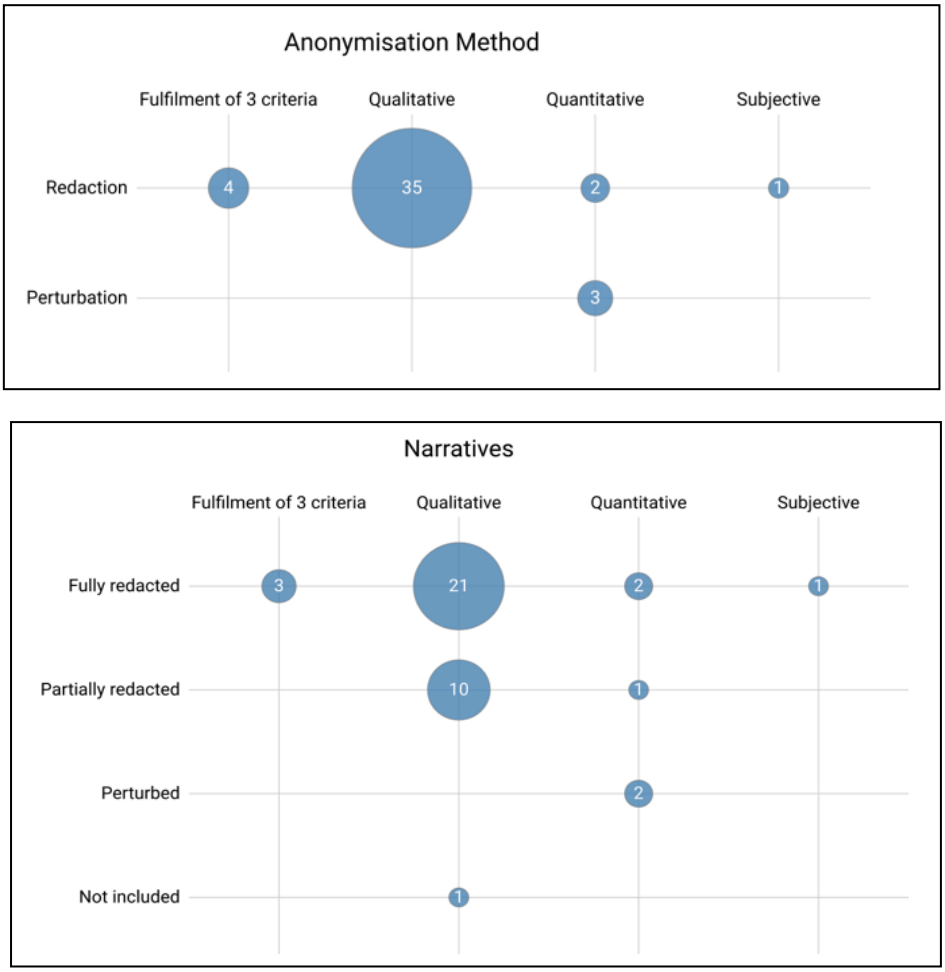

Figure 1 - Summary of Findings on first 57 Policy 0070 Submissions

EMA released a report over the same period *"Clinical Data Publication report oct16-oct17"* [5] where more details on how confidential information was assessed, accepted and rejected was shared. The conclusions therein related to de-identification techniques are similar than the ones of the PhUSE Data Transparency Working Group.

It became apparent that a need for further guidance beyond the scope of the Policy 0070 External Guidance was required to help sponsors meet the Policy 0070 anonymization preferred approach (i.e. Quantitative and Anonymisation) and data utility objectives.

EMA TECHNICAL ANONYMISATION GROUP

In addition, and as a result of the first set of Policy 0070 submissions, the EMA formed a Technical Anonymisation Group (TAG), [6] with individuals representing different stakeholders and skill sets, with the following objective:

The European Medicines Agency (EMA) has established an expert group in data anonymisation known as the technical anonymisation group (TAG) to help further develop best practices for the anonymisation of clinical reports, in the context of the Agency's policy on the publication of clinical data.

The group includes members from academia, industry, patients and healthcare professionals with expertise in areas such as data protection, developing standards and guidance for anonymisation and re-analysis of clinical data. EMA established the group following a public call for applications launched in March 2017.

The group will consider experience to date with EMA's publication of clinical reports. In particular, it will assess:

- *patient re-identification and any privacy risks in the light of new technological developments;*
- *the scientific utility of the published clinical data as a function of the anonymisation methodology used;*
- *whether it is possible to successfully conduct a secondary analysis of the anonymised clinical data.*

The application process consisted of sending a CV and a Declaration of Interest (DoI). Among 45 applications received, the following 14 individuals were selected:

Table 2 - EMA TAG Members, affiliations and countries

Pharmaceuticals R&D / Data Sharing (2)	**Academia – Medical Science (3)**
Christine Fletcher (Amgen) - UK	Sarah Nevitt (University of Liverpool) - UK
Lukasz Kniola (Biogen) - UK	Frank Rockhold (Duke University) - USA
Pharmaceuticals / CRO Legal (3)	Kristian Svendsen (Hospital Pharmacy Tromsø) - Norway
Nicola Orlandi (Novartis) - Switzerland	**Software (2)**
Lee Parker (Biogen) - UK	Khaled El Emam (Privacy Analytics) - Canada
Uwe Fiedler (PAREXEL) - Germany	Cathal Gallagher (d-Wise) - UK
Patient Association (1)	**SME (2)**
Rafal Swierzewski (EPPOSI) - Belgium	Jean-Marc Ferran (Qualiance) - Denmark
Data Protection Agency (1)	Bradley Malin (Vanderbilt University) - USA
Guiseppe d'Acquisto (Italian DPA) - Italy	

In addition, a member of European Data Protection Supervisor (Dina Kampouraki) participates as an observer.

EMA TAG Members are contributing to five topics:

- Anonymisation Techniques
- Data Utility
- Legal Issues/GDPR
- New Technological developments
- Potential Attackers

Sub-teams are meeting remotely and members draft deliverables. One Face-to-Face meeting takes place at the agency with all members. The TAG will develop additional guidance (e.g. Q&A) to further clarify certain aspects of the methodology described in the external guidance.

HEALTH CANADA'S DOCUMENTS FOR PUBLIC REVIEW

Health Canada launched their Data Transparency initiative in 2017 and provided the opportunity to review the following documents each time with a 75-day timeline:

Table 3 - Health Canada's Data Transparency Documents for Public Review

Document*	**Release Date for Public Review**	**Purpose**
White Paper [7]	10MAR2017	*"This document introduces Health Canada's initiative to publicly release clinical information concerning the safety and efficacy of drugs and safety and effectiveness of medical devices. It sets out the policy objectives, rationale and considerations for future regulations that would specify that certain clinical information contained in drug submissions or medical device applications would not be treated or cease to be confidential business information following a final regulatory decision and that would authorize the public release of that information."*
Regulation [8]	09DEC2017	*"The objective of this proposal is to provide public access to clinical information submitted to Health Canada in drug submissions for human use and medical device applications following a final regulatory decision…This proposal contains amendments to the Food and Drug Regulations and the Medical Devices Regulations. Proposed amendments to the Food and Drug Regulations would specify the kind of clinical information in drug submissions that would cease to be CBI following a final regulatory decision."*
Guidance [9]	10APR2018	*"This document is designed to help the public, industry, healthcare professionals and other stakeholders better understand the implementation of Health Canada's Public Release of Clinical Information initiative, including: the procedures to prepare information for release; the categories of information that continue to be subject to the definition of confidential business information (CBI) and that may*

		be eligible for redaction; and how Health Canada will protect personal information."

* The White Paper was restricted for comments to Canadians or Canadian organizations only (e.g. the affiliate of a given sponsor could send their comments), while there was not such restriction on the two other documents.

Health Canada provided an anonymized summary [6] of the results of the consultation on the White Paper. Figure 2 represents the distribution of unique submissions from each stakeholder group (some submissions were submitted on behalf of multiple individuals or organizations).

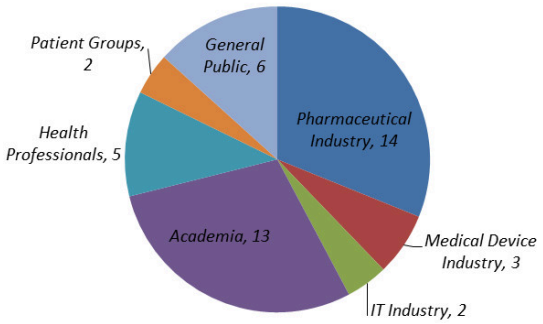

Figure 2 - Submissions in response to Health Canada's White Paper by stakeholder groups

The pharmaceutical industry and academia were the two largest contributors.

In particular, pharmaceutical industry's comments were encouraging Health Canada to align as much as possible with Policy 0070 requirements while also highlighting current challenges with Policy 0070. Additionally, comments highlighted that conversations with EMA were also progressing. Among other comments, health professionals raised *"concerns with respect to risks of delay in implementation and proper resourcing of this initiative."* and academia noted that *"consultation on the scope of the redactions and time allowed for industry should be well defined in order not to prolong the process."*. A complete summary is available at [10].

The final regulation and guidance was anticipated to be published at the end of the year 2018 as announced at DIA in London in September 2018, however current expectations are for publication in early 2019.

The main points of Health Canada's Draft Guidance are:

- Health Canada validates anonymisation approach
- Only Quantitative approach is mentioned for residual risk analysis
- Recognition process with EMA Policy 0070 documents is possible through a certification process
- Four Reference Populations are listed to consider in modeling the risk and compute the residual risk
- It is possible to request access to past submissions documents following a prioritization scheme

HEALTH CANADA'S STAKEHOLDERS GROUP FOR PUBLIC RELEASE OF CLINICAL DATA

Following the public review of their White Paper, Health Canada contacted stakeholders to apply to their Stakeholders Group:

In follow up to our public consultation, [Health Products and Food Branch] HPFB will engage with experts and stakeholders on implementation of public release of clinical information. To support this process HPFB intends to establish a roster of experts who are available to provide advice, specific to the Canadian context, in the following areas:

- *Clinical trial reporting, medical writing*
- *Regulatory affairs & CTD structure*
- *Health data science and statistics*
- *Data protection and privacy law*
- *Anonymization and de-identification methods*
- *Open data and end-user experience*

An application letter and a DoI were requested to apply. The following 15 members were subsequently selected:

Table 3 - Health Canada Stakeholders Group's Members, affiliations and countries

Pharmaceuticals Regulatory Affairs (4)	**Academia - Law & Policy (2)**
Arshia Ghani (Pfizer) - Canada	Matthew Herder (Health Law Institute, Dalhousie University) – Canada
Noel Geaves (Medtronic) - Canada	Joel Lexchin (School of Health Policy & Management, York University) – Canada
Sarah Marshall (NAPRA) – Canada	**Academia – Medical Science (2)**
Sandra Wainwright (Merck) - Canada	Nav Persaud (Family & Community Medicine, University of Toronto) – Canada
Sally Prawdzik (J&J) - Canada	Angela Spelsberg (Comprehensive Cancer Center Transparency International) – Germany
Pharmaceuticals R&D / Data Sharing (2)	**Software (1)**
Milan Shah (J&J) - USA	Hazel Nicholls (Privacy Analytics) - Canada
Teresa Armstrong (Eli Lilly) - USA	**Standards Organization (1)**
Professional Organization (2)	Jean-Marc Ferran (PhUSE) – Denmark
Barry Power (Canadian Pharmacists Association) - Canada	
Bradley Mitchelmore (Canadian Agency for Drugs and Technologies in Health) - Canada	

The Stakeholders' Group met at the pace of one 2-hour teleconference per month where documents drafted by Health Canada's officers were sent in advance and discussed during each call. The themes were:

- Kick-Off Meeting and review of overall draft process (October 2017)
- Overview of operations and key implementation issues (November 2017)
- Protection of personal information (January 2018)
- Process to specify clinical information, safeguards against commercial use (February 2018)
- Public release options and health system impacts (April 2018)

Health Canada officers used the comments and feedback from the Stakeholders' Group to further elaborate some aspects of the Draft Guidance. The Draft Guidance was discussed with the Stakeholder Group after its initial publication for public review.

FDA DATA TRANSPARENCY INITIATIVE AND PILOTS WITH SPONSORS

FDA launched their Clinical Data Summary Pilot Program [11] in January 2018. FDA has extensive experience in redacting clinical documents through their processing of Freedom of Information Act (FOIA) requests since 1996.

Objectives refer to *"making a CSR publicly available after a drug's approval will provide stakeholders with more information on the clinical evidence supporting a drug application and more transparency into the FDA's decision-making process."*. The press release does not refer to secondary analysis or more specific goals but only to a general data transparency principle related to the agency decision making process.

Main points from FDA initiative:

- FDA is "doing the work" and redacts the documents. Sponsor is not consulted but "*If the sponsor is uncertain about whether we will redact certain types of information, the sponsor should ask us (i.e. FDA) in advance"*
- Only redaction as a de-identification method is listed. In particular "*Demographic information, such as sex, age, and race, will generally **not** be redacted, except in very unusual circumstances"*
- Patient/subject IDs, as well as any other unique patient identifier (e.g., Social Security number), and patient contact information will be redacted;
- Signatures, including wet signatures, and business contact information will **not** be redacted when the individual's name (e.g., name of a sponsor's employee) is otherwise disclosable.
- Narratives are out of scope
- Only pivotal phase III studies are in scope

JAPAN – LAW CONCERNING ACCESS TO INFORMATION HELD BY ADMINISTRATIVE ORGANIZATIONS

Information on Japanese Regulatory Affairs [12] describes that,

> "A notification concerning publication of information on new drug approvals was issued (Notification No.1651 of the Evaluation and Licensing Division, PMSB [Pharmaceutical and Medical Safety Bureau] dated November

> 11, 1999), and New Drug Approval Information Packages containing summary reviews prepared by the MHLW [Ministry of Health, Labor and Welfare] and nonclinical and clinical data submitted by the applicant have been published."

This document further describes that,

> "With the enactment of the Law Concerning Access to Information Held by Administrative Organizations on April 1, 2001, anyone has the right to request disclosure of documents retained by national government organizations. This law covers disclosure of documents retained by government organizations except those concerning non-disclosable information such as information on individuals, information on corporations, etc. This was partially amended by Cabinet Order No. 371, on December 21, 2005.
>
> Based on this Law, the MHLW must disclose the contents of its reviews (records of meetings of the PAFSC [Pharmaceutical Affairs and Sanitation Council], new drug approval information dossiers, etc.), as a rule, and new procedures for processing work related to public disclosure of information retained by the PFSB [Pharmaceutical and Food Safety Bureau] were specified (Notification No. 0330022 of the PFSB dated March 30, 2007)."

On 25 March 2013, a Partial Revision of the Disclosure of Information on Approval Evaluation of New Medicinal Products (PFSB/ELD Notification No.0325-1) [13] was issued, and further operational requirements are described therein.

These documents describe requirements related to the publication of common technical documents (CTDs) and the masking of personal information and confidential information.

Documents in-scope for disclosure are classified into six types:

- evaluation and licensing-related documents,
- safety-related documents,
- compliance-related documents,
- narcotics-related documents,
- blood and blood products-related documents, and
- other activity-related documents.

and documents are marked as ○ (disclosure), ● (non-disclosure) or Δ (partial disclosure).

Related to approval application documentation, which is generally not accessible until after approval, the following sections are considered in scope for disclosure:

- Review report including module 1.15-1.10 and 1.12 (note this is different to EU module 1),
- All Module 2 (Plus additional requirements including Clinical Study Report (CSR) synopses and mini-narratives for serious adverse events (SAEs)).

Information in these sections is redacted if related to individuals or to company-confidential information.

Conversely, Module 3 ("Quality-Related Documentation"), Module 4 ("Nonclinical Study Reports"), and Module 5 ("Clinical Study Reports") are not in scope for disclosure.

SUMMARY

The table below summarize how the four agencies have engaged with Stakeholders on their Data Transparency initiatives. All Agencies have actively participated in conferences and webinars to inform stakeholders and discuss comments and concerns. The table below describes formal engagement only.

Table 4 - Overview of differences and similarities between agencies on engaging Stakeholders

Item	EMA Policy 0070	Health Canada Public Release of Clinical Information	FDA Clinical Data Summary Pilot Program	PMDA Disclosure of Information
Consultation on:	Policy & Guidance through Stakeholders Review	Policy & Guidance through Public Review	Planned *"public feedback through a Federal Register notice and docket for public comments"* following the conduct of the pilots	NA
Pilot	None	None	Up to nine pilots planned with sponsors	None

Item	EMA Policy 0070	Health Canada Public Release of Clinical Information	FDA Clinical Data Summary Pilot Program	PMDA Disclosure of Information
Working Groups	EMA TAG	Stakeholders Group	Not planned so far	NA
Working Groups Application	Public call for applications CV & DoI required	Individual call for applications Application Letter & DoI required	NA	NA
Working Group Mandate	Maximum 2 years renewable	6 months between October 2017 and April 2018	NA	NA
Working Groups Deliverables	Q&A, Additional Guidance, Critical Review (TBA) developed by TAG members under EMA officers' supervisions	Participation in five Meetings to comment on five key topics that led to the development of the guidance by Health Canada officers.	NA	NA

NA – Not Available

FDA is the only agency that has taken an approach involving pilots that would help inform the final policy. These pilots are beyond just prototypes performed "behind closed doors". One pilot performed with J&J (Janssen Pharmaceuticals) has been published so far, and redacted submission documents including CSRs, protocols and Statistical Analysis Plans (SAPs) is available for public access with no geographical restriction [14].

EMA and Health Canada approaches are similar in terms of involving stakeholders in reviewing policy and guidance and running Working Groups to further advise them. Their two Working Groups are different in two aspects:

- Timelines (2 years vs. 5 months) and Involvement (Drafting Deliverables vs. Commenting)
- Working Groups Members: Health Canada had many professionals coming from regulatory groups within pharmaceutical companies, while EMA had a more diverse representation also involving lawyers and R&D staff from pharmaceutical companies.

The second difference related to membership can be explained by the timing of the initiatives and the fact that EMA had already released a guidance and was running their initiative while Health Canada was still scoping their guidance and could benefit from EMA's experience.

EMA also refers in their external guidance ([3] Chapter 2, section 5.1 Data Utility) to a qualitative approach and use of redaction in an initial phase until sponsors gain more experience and can use quantitative approaches and anonymization. There is no pilot as such, but the guidance suggests that there is a period where sponsors can develop their skills to transition to quantitative and anonymization techniques. Health Canada Draft Guidance does not refer to such "grace period". In contrast, FDA applies a redaction approach as they use to also fulfil FOIA requests, whereas in Japan, the PMDA approach largely requires redaction with some anonymization techniques also applied.

REQUIREMENTS COMPARISON

The four initiatives from the agencies differ in terms of:

- Terminology
- Scope and Processes
- Submission documentDe-identification Requirements

TERMINOLOGY

When reviewing the different documents, one can notice subtle differences in terminology. Table summarizes differences on main terms.

Table 5 - Overview of differences in Terminology

EMA Policy 0070 External Guidance	Health Canada Public Release of Clinical Information Draft Guidance	FDA Clinical Data Summary Pilot Program	PMDA Disclosure of Information
Anonymisation De-Identification Transformation	Anonymization De-identification Transformation	Redaction	Redaction, with limited anonymization techniques e.g., rounding of age and generalization of dates
Direct Identifers	Directly-identifying variable	Unique Patient Identifier	Unique Patient Identifier Site/investigator ID may be proposed as CCI redaction
Quasi Identifiers	Indirectly-identifying variable	No equivalent term although such identifiers are discussed but not named as such	No equivalent term
Protected Personal Data (PPD)	Personal Information	Personal Privacy Information	NA though some redactions of subject ID have been agreed
Data Utility	Analytical Value/Analytical Utility Data Utility	No equivalent term, concept not addressed	NA
Risk of Re-identification Residual Risk	Risk of Re-identification Data Risk	NA	NA
Commercially Confidential Information (CCI)	Confidential Business Information (CBI)	Confidential Commercial Information Trade secret	Information which is considered to harm the competitive stance or other legitimate interest

NA - Not Available

Certain differences are driven by design i.e., by the different approaches taken by the agencies e.g. redaction for FDA vs. Anonymisation/De-Identification for EMA and Health Canada, and chiefly redaction for PMDA. It is not always clear from the different documents whether other terms hold any differences in meaning (e.g., Commercially Confidential Information vs. Confidential Business Information vs Confidential Commercial Information, Data Utility vs. Analytical Value) and some alignment could contribute to provide clarity to the different stakeholders as the policies cover many of the same documents. Certain documents also interchangeably refer to different terms for the same concepts (Anonymization/Anonymisation and De-Identification, Risk of Re-identification and Residual Risk/Data Risk), and efforts would be welcomed in future versions to align intra- and inter-documents across agencies, where possible.

SCOPE & PROCESS

The four agencies differ in scope and processes as highlighted in table below.

Table 6 - Overview of differences in Scope & Processes

Item	EMA Policy 0070 External Guidance	Health Canada Public Release of Clinical Information Draft Guidance	FDA Clinical Data Summary Pilot Program	PMDA Disclosure of Information
Effective Date (submission from)	MAAs or part of 'Article 58' procedures from 01 January 2015 and extensions of indications from 01 July 2015 –	Not effective yet	Not effective yet	11 November 1999
Retrospective access to past documents	Not in scope Possibility through	In scope, based on prioritization system	Not in scope	Not in scope, though as the

Item	EMA Policy 0070 External Guidance	Health Canada Public Release of Clinical Information Draft Guidance	FDA Clinical Data Summary Pilot Program	PMDA Disclosure of Information
	Policy 0043. Sponsors review the redactions proposed by EMA. Redacted documents are sent to requester.	using *metrics that identify products and information with high health system impact. Sponsors conduct anonymization unless a certified EMA Policy 0070 document is available. Documents are made public.*	Possibility through FOIA request FDA conduct the redactions. Documents are sent to requester.	effective date is from 1999, many submissions are in scope of the proactive process
Studies in Scope	All studies part of a Central Application regardless of submission outcome[1]	Step-wise approach over 4 years including Medical Devices studies from year 3. Regardless of submission outcome	Only Phase III pivotal studies CSRs following approval of an NDA	CSR synopses included in Module 2 are in scope, but full CSRs in Module 5 are out of scope
Timelines to submit de-identified documents to agency	<= 30 days pre-opinion and <= 10 days post-opinion. NOTE: Until now, EMA has contacted sponsors to address first "back-log" of submissions rather than following these timelines.	20 days after notification	Not discussed	Draft redactions to be submitted by sponsor around approval date, and will be published within a month of approval
Recognition Process with other Agencies	None	Yes, with EMA through a certification application	Not discussed	None
Review Process (de-identification of PPD)	Using Annotated Documents and Anonymisation Report EMA provides recommendation and comments	Same as EMA but HC validates de-identification of patient information and keeps decisions on what is publicly released.	NA but Sponsor can notify FDA of special-attention items in CSRs	Sponsor states and substantiates rationale for redaction PMDA reviews and has final decision
Handling of CCI/CBI	5-rejection criteria	Two categories: • Clinical information that was not used by the manufacturer in the submission • Clinical information that describes tests, methods or assays used exclusively by the manufacturer	Confidential Commercial Information and Trade Secrets as defined in the context of FOIA are redacted.	Information considered to harm the competitive stance or other legitimate interests of the corporation may be redacted.

[1] "*Where clinical study reports are cross-referred to within paediatric extension or modification of indication applications, the MAH is required to submit for publication pivotal clinical study reports as well as all supportive studies conducted in the paediatric population that were submitted in the context of regulatory procedures not falling within the scope of Policy 0070 and considered the basis for that application.*"

Item	EMA Policy 0070 External Guidance	Health Canada Public Release of Clinical Information Draft Guidance	FDA Clinical Data Summary Pilot Program	PMDA Disclosure of Information
IPD in scope	Planned in Phase 2	Could be in the future	Not in scope of present initiative	No
Who can access the report?	Users registered as EU citizens. Non-citizens can view but cannot download reports.	Users registered on Health Canada website	Anyone without registration	Publicly available, though mostly available in Japanese
Term of Use	Yes	Yes	None	None

One of the most notable differences is that FDA's approach only considers pivotal Phase III studies as in scope following approval of an NDA, while EMA and Health Canada do not set a condition related to submission outcomes. PMDA make documents available after approval, but their scope of document release is more in line with that of EMA and Health Canada (though PMDA releases CSR synopses not full CSRs for in-scope studies).

Another important difference is related to the review process of the de-identification of PPD. It is not possible in the case of the FDA approach for the sponsor to directly participate in the process, *"the sponsor will not have an opportunity to review our redactions before the CSR is posted. If the sponsor is uncertain about whether we will redact certain types of information, the sponsor should ask us in advance"*. In the case of EMA Policy 0070, sponsors drive the de-identification process, define the anonymization/redaction level and discuss comments and recommendations from EMA Policy 0070 review teams during the submission process. In the latest version of the external guidance (version 1.4, 15 October 2018), sponsors must confirm in a check-list that *"Anonymisation report has been revised in line with the Agency's recommendations / comments"*. As per Health Canada Draft Guidance, the anonymization level must be agreed with the Agency upfront, meaning that the agency would share a greater role as Data Controller. For submissions in Japan, the sponsor needs to state and substantiate the rationale for each redaction, PMDA will review this rationale and has the final decision whether the data item is redacted or not.

CSR DE-IDENTIFICATION REQUIREMENTS

De-identification requirements are also different across the four agencies as highlighted in table below.

Table 7 - Overview of differences in CSR De-identification Requirements

Item	EMA Policy 0070 External Guidance	Health Canada Public Release of Clinical Information Draft Guidance	FDA Clinical Data Summary Pilot Program	PMDA Disclosure of Information
Narratives in scope	Yes	Yes	No	Module 2 documents include CSR synopses and mini-narratives for all SAEs
Risk analysis	Working Party 29 Opinion 3-Criteria Qualitative Quantitative	Quantitative only	Qualitative based on FDA's approach for FOIA request	Qualitative – sponsor to justify all proposed redactions
De-identification Technique	Anonymization & Redaction	Anonymization	Redaction only based on FDA's approach for FOIA request. In particular, "*Demographic information, such as sex, age, and race, will generally **not** be redacted, except in very unusual circumstances"*	Largely redaction with some transformations. PPD redaction is limited.
Guidance on Identification of	Refer to PhUSE Standard [15] and discuss in particular	Indirectly-identifying variables are other identifying variables	Different type of identifiers discussed in Q&A page [11]:	The 2013 notification (No. 0325) provides

Item	EMA Policy 0070 External Guidance	Health Canada Public Release of Clinical Information Draft Guidance	FDA Clinical Data Summary Pilot Program	PMDA Disclosure of Information
Direct/Quasi Identifiers	quasi identifiers such as: Dates, Geographic Location, Other Quasi-Identifiers (e.g. Demographics)	that fall within the definition of 'personal information' within Canada's Privacy Act. And refer to Demographics and medical history and SAE. ***Country should remain unmodified.*** HIPS* is encouraged	Unique Patient Identifiers Dates Clinical Trial Site Geographic Location Demographics Relative dates (study days) are retained [14]	some guidance: Redaction, with limited anonymization techniques e.g., rounding of age and generalization of dates
Guidance on Reference Population	No direct guidance but examples of plausible attacks to consider for risk modeling are listed.	Four populations are provided for consideration: • Study Population • Similar Sponsor Trials Population • Similar Trials Population • General Geographic Population	NA	NA
Risk Threshold	Default 0.09, possible to justify another one	Default 0.09, possible to justify another one	NA	NA
Anonymisation Report	Yes	Yes, similar to EMA	NA	No, though all proposed redactions must be justified by the sponsor to PMDA
Guidance on Data Utility	No definition provided Must be justified in Anonymization report Refer to not altering interpretation of study results and ability to produce similar analysis results	No definition provided Must be justified in Anonymization report as to *state the efforts made to maximize the utility of the anonymized information*.	Not Addressed	Not Addressed
Personal Data of Investigators, Sponsor staff and Providers	Only sponsor and coordinating investigator signatories of the CSR are kept	Not Addressed so far	*Generally **not** redacted if the contract relationship is public*	*Investigator names and the names of CROs are redacted if not already in the public domain (redacted as CCI)*

* HIPS: Hiding in Plain Sight; NA – Not Available

While the EMA and Health Canada requirements are similar, the FDA requirements are very different starting with the "redaction-only" de-identification technique, and the fact that Demographics *will generally not be redacted*, which is not

possible to apply when using quantitative methods as advised by the two other agencies. It must also be noted that while EMA and Health Canada emphasize the importance of data utility (or analytical value) and justifying how sponsors have optimized it, neither have provided a clear definition or tangible guidance on how to do so. Also noteworthy is that narratives are out-of-scope of the FDA Data Transparency pilot.

FDA retained relative days (study day, e.g. Study Day 126) in their first pilot [14] and redacted actual dates (e.g. 12 April 2015) as illustrated in figure below. Sponsors could consider using relative study days rather than actual dates when developing CSRs in future, as relative study day is also considered anonymized under EMA Policy 0070 Guidance and Health Canada Draft Guidance. The only weakness with study day is that it is not possible though to represent partial dates (e.g. April 2015).

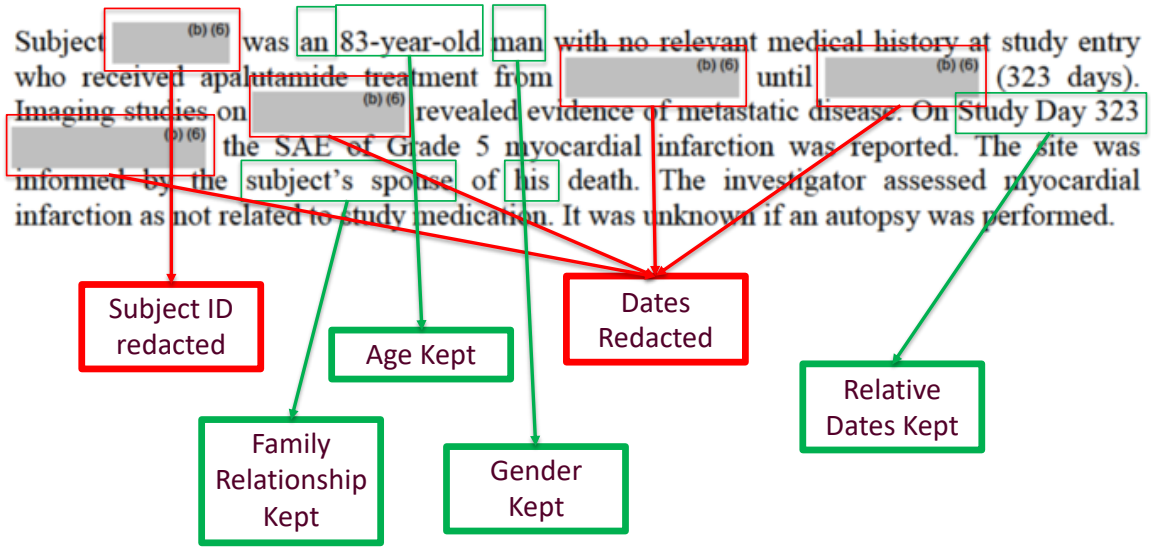

Figure 3 - Example from FDA Pilot (ERLEADA) [14]

In Figure 3, it must be noted that Age (here "83-year-old") is not redacted in ERLEADA pilot. Under EMA Policy 0070 and Health Canada Draft Guidance, Age would likely be generalized to age intervals or shuffled within an age interval to meet the 0.09 maximum risk threshold. Assuming the same CSRs are published in different jurisdictions according to different requirements, an adversary could use exact age from the FDA version of the redacted CSRs in connection with other anonymized data in CSRs published through EMA Policy 0070 or Health Canada's initiative.

Health Canada's Draft Guidance refers to four different reference populations from which to select when modelling the risk and plausible adversaries, and they emphasize that the study population (the smallest one) must not be selected by default. This point was not highlighted in the EMA Policy 0070 guidance.

Health Canada's Draft Guidance advises that *Country should remain unmodified* [9]. This is not in alignment with accepted de-identification practices where country does not always hold a high data utility while being one of the first variables to generalize in order to enhance the population when considering risk calculations, as advised by the PhUSE standard [15].

Based on some sponsors' experience in Japan, dates such as submission dates and data cut-off dates can often be redacted as CCI. Regarding dates relating to subjects, for key dates such as SAE start dates, the YEAR is redacted and replaced with a code, e.g., 'within first year of treatment' but the MONTH and DAY are retained. For less critical dates, the YEAR is often redacted. Given additional sensitivities around certain dates such as date of birth/death, there can be more extensive redaction or data transformation. Age may be partially redacted, for example, generalized into 10-year age bands.

DISCUSSION

At the time this paper is written, EMA has announced to pause Policy 0070 activities after the publication of documents from 142 submissions, until after their relocation to Amsterdam, while Health Canada is targeting to release policy and guidance in early 2019. Only one submission has been published under the FDA's pilot, and the process in Japan has been in effect for almost 20 years. During this time, the Japanese process has published approximately 60 dossiers, for which the information related to 4 HIV products has been released in English and information from the other dossiers is available in Japanese. The conclusions below are extrapolative and may need to be adjusted or confirmed once some agencies have progressed further with their initiatives.

EMA's and Health Canada's initiatives are comparable while FDA's differs in many ways. In particular, FDA only uses redaction and processes documents for personal information and CCI/trade secrets on behalf of the sponsors. While de-identified documents processed through EMA's and Health Canada's initiatives will be similar, important differences with FDA-processed documents are to be expected. In particular, narratives are not in scope of FDA data transparency program while demographics are not necessarily redacted. EMA's and Health Canada's initiatives mainly differ on the approach to risk calculation where only a quantitative approach is considered under Health Canada, and Health Canada is also validating the sponsors' anonymization approaches. In addition, Health Canada will provide access to documents from past submissions ("retrospective requests") following a prioritization scheme. Publication of (sponsor) signatures is also a difference seen between regions. Whereas EMA and Health Canada processes will expect signatures to be redacted, signatures are required to be disclosed under the FDA pilot. In Japan, signatures are not included in either module 1 or 2.

Sponsors will have to adjust to these various policies, and process documents such that consideration is given as to what may be released at a later stage by another agency in another jurisdiction. In particular, the FDA plans to conduct the redaction without consulting the sponsors, which would mean that sponsors need to understand very well FDA's redaction approach and to take it in consideration when de-identifying documents for EMA or Health Canada. conversely, FDA could consider other de-identified documents published/to be published in other jurisdictions when conducting their redactions of pivotal Phase III study CSRs that are part of a positive NDA. In particular, the retention of age and other demographics by the FDA could be problematic and invalidate the residual risk analysis of de-identified documents otherwise published through EMA and Health Canada.

A number of sponsors are currently analyzing how CSRs could be written at the time the study is reported, in a way that facilitates anonymization/redaction later on across agencies at the time of submissions or approvals. The use of relative days such as study day (instead of actual dates to be offset or redacted) appears to be a good practice as such a format is considered anonymized under EMA's and Health Canada's approach, and is retained also by the FDA [14] following their FOIA-like approach and by Japan. At the time this paper is written, Health Canada's Draft Guidance advises that *Country should remain unmodified* [9], which is not in alignment with accepted de-identification practices [15]. Should this requirement be confirmed in the final guidance, sponsors might consider under EMA Policy 0070 as a result, to retain country if at all possible and potentially anonymize/redact further other variables than those considered necessary for EMA, so as to consider a more global approach to the disclosure of information, even though this may decrease clinical data utility.

When the different policies are all effective, it will be interesting to analyze which of the different sources will be used by the different data consumers [16] and which documents will hold the most data utility, in particular for the academic research community. Since discussions around sharing IPD are postponed by all agencies, the main source for IPD remain through sponsors' own voluntary Data Sharing initiatives.

The possibility that EMA documents are accepted for Health Canada submissions through a certification process is positive news for the industry and would help sponsors save resources should the two agencies keep their requirements rather close to each other. More importantly, this could also have a positive impact on the level of clinical data utility within the published documents as sponsors would be able to focus more on data utility rather than finding the de-identification "minimum common denominator" across agencies' requirements. It is not clear yet from the Health Canada Draft Guidance how critical the certification process from Health Canada will be though, and whether alignment with Health Canada's de-identification requirements will be considered to provide certifications. Considering the pause of EMA Policy 0070, it will also be interesting to see whether EMA will envisage a similar certification process for documents published first under Health Canada's process.

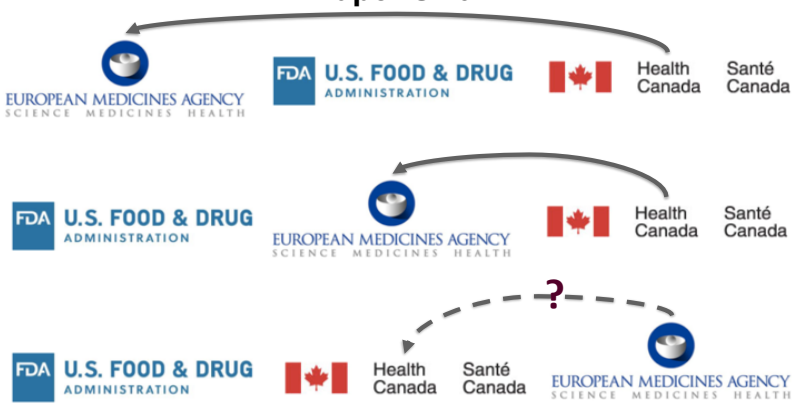

Figure 4 - Recognition/Certification Process between Agencies?

It must also be noted that if de-identification requirements are not aligned and recognition/certification processes not critical, it could create a gap in clinical data transparency where the publication of a given document in the first jurisdiction would dictate the level of de-identification and data utility.

EMA, Health Canada and FDA all provide access to CSRs submitted to the agencies prior to the effective date - through EMA Policy 0043 (now available to be requested by EU residents only) and FOIA requests for FDA. For EMA Policy 0043, sponsors prepare redactions while FDA conduct redactions for FOIA requests. In both cases, redacted documents are sent to requesters only and not posted publicly. Health Canada's approach includes the possibility to submit requests for documents from previously approved products, where sponsors would conduct anonymization steps (not redaction only) following the approach described in the Draft Guidance and involving validation from Health Canada. The anonymized documents would then be made public. There is no limit to how far back in time documents can be requested, which may lead to technical challenges for sponsors to anonymize some legacy CSRs, rather than redacting them. In Japan, requirements to publicly post information has been implemented for almost 20 years and mechanisms for retrospective requests are not covered in this approach.

The four agencies must also address the question of Data Controllership and whether a joint controllership is to be considered, as well as clarification of responsibilities. While EMA provide comments and recommendations based on review of anonymization reports, Health Canada validates anonymization approach and FDA conducts redaction without consulting sponsors. In revision to Japanese guidance documents in March 2013, it was clarified that, "'the applicant retains the rights on information in the data and the accountabilities they contain". There are clearly at this stage varying levels of Data Controllership where EMA has the least responsibility and FDA has the greatest. Japan and Health Canada have an intermediate level of responsibility in case of data breaches.

REFERENCES

[0] Documents from Advisory Groups on Clinical Trial Data:
https://www.ema.europa.eu/en/human-regulatory/marketing-authorisation/clinical-data-publication/documents-advisory-groups-clinical-trial-data
[1] EMA initial policy of 24. June 2013:
https://www.ema.europa.eu/documents/other/draft-policy-70-publication-access-clinical-trial-data_en.pdf
[2] EMA Policy 0070 of 2nd October 2014:
https://www.ema.europa.eu/documents/other/european-medicines-agency-policy-publication-clinical-data-medicinal-products-human-use_en.pdf
[3] EMA Policy 0070 External Guidance versions:
https://www.ema.europa.eu/human-regulatory/marketing-authorisation/clinical-data-publication/support-industry/external-guidance-implementation-european-medicines-agency-policy-publication-clinical-data
[4] Analysis of CSRs already published, PhUSE Review – Lukasz Kniola, November 2017
https://www.ema.europa.eu/documents/presentation/presentation-phuses-analysis-clinical-reports-published-l-kniola_en.pdf
[5] EMA Clinical Data Publication (Policy 0070) report Oct 2016-2017
https://www.ema.europa.eu/documents/report/clinical-data-publication-policy-0070-report-oct-2016-oct-2017_en.pdf
[6] EMA TAG page:
https://www.ema.europa.eu/human-regulatory/marketing-authorisation/clinical-data-publication/technical-anonymisation-group
[7] Health Canada's White Paper
https://www.canada.ca/en/health-canada/programs/public-release-clinical-information-drug-submissions-medical-device-applications.html
[8] Health Canada's Regulation
http://www.gazette.gc.ca/rp-pr/p1/2017/2017-12-09/pdf/g1-15149.pdf
[9] Health Canada's Draft Guidance:
https://www.canada.ca/en/health-canada/programs/consultation-public-release-clinical-information-drug-submissions-medical-device-applications/draft-guidance.html
[10] Health Canada's Responses to the Public Consultation on the White Paper "Public Release of Clinical Information in Drug Submissions and Medical Device Applications":
https://www.canada.ca/en/health-canada/programs/consultation-public-release-clinical-information-drug-submissions-medical-device-applications/what-we-heard.html
[11] FDA Clinical Data Summary Program page:
https://www.fda.gov/drugs/developmentapprovalprocess/ucm589210.htm
[12] INFORMATION ON JAPANESE REGULATORY AFFAIRS 2018
http://www.jpma.or.jp/english/parj/pdf/2018.pdf
[13] Partial Revision of the Disclosure of Information on Approval Evaluation of New Medicinal Products, PFSB/ELD Notification No.0325-1, March 25, 2013.
[14] "Drug Approval Package: ERLEADA" posted on FDA website:
https://www.accessdata.fda.gov/drugsatfda_docs/nda/2018/Erleada_210951_toc.cfm
[15] PhUSE Data De-Identification Standard for SDTM 3.2 (2015)
http://www.phuse.eu/data-transparency-download
[16] Ferran, Nevitt 2017, EMA Policy 0070: Data Utility in Anonymised Clinical Study Reports (CSRs), DH04 - PhUSE Annual Conference 2017
https://www.phusewiki.org/docs/Conference%202017%20DH%20Papers/DH04%20Paper%20NEW.pdf

GLOSSARY

Term	Definition
Anonymization	Anonymization refers to different transformation techniques such as generalization, permutation, perturbation, date off-setting etc. as opposed to redaction only
CBI	Confidential Business Information
CCI	Commercially Confidential Infirmation
CRO	Contract Research Organization
CSR	Clinical Study Report
CTAG	Clinical Trial Advisory Groups - EMA cross-industry working groups established in 2012 to develop a clinical data publication policy
CTDs	Common Technical Documents
De-Identification	De-identification refers generally in this document to anonymization or redaction techniques
Direct Identifiers	One or more direct identifiers can be used to uniquely identify an individual. E.g. Subject ID, Social Security Number, Telephone number, Exact address, etc. It is compulsory to remove or de-identify any direct identifier.
DoI	Declaration of Interests
EMA	European Medicines Agency
FDA	Food and Drug Administration - US Agency
FOIA	Freedom Of Information Act
HCRG	Refers to Health Canada Reference Group for public disclosure or clinical information
HIPS	Hiding in Plain Sight - Method consisting in replacing the detected identifiers with its anonymized version in text in order to make 'leaked' identifiers difficult to distinguish.
HPFB	Health Products and Food Branch - Health Canada division
IPD	Individual Patients Data
MAA	Marketing Authorization Application
MAH	Marketing Authorization Holder
MHLW	Ministry of Health, Labor and Welfare (Japan)
NDA	New Drug Application (FDA)
PAFSC	Pharmaceutical Affairs and Sanitation Council (Japan)
PFSB	Pharmaceutical and Food Safety Bureau (Japan)
PMDA	Pharmaceutical and Medical Devices Agency - Japanese Agency
PMSB	Pharmaceutical and Medical Safety Bureau (Japan)
PPD	Protected Personal Data
Quasi Identifiers	Quasi identifiers are background information that can be used in connection with other information to identify an individual with a high probability. E.g. Age at baseline, Race, Sex, Events, Specific Findings, etc.
Redaction	Redaction refers to the blacking-out of portions of text
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
TAG	Refers to EMA Technical Anonymization Group
Working Party 29	Refers to Article 29 Working Party - The Article 29 Working Party was an advisory body made up of a representative from the data protection authority of each EU Member State, the European Data Protection Supervisor and the European Commission.

ACKNOWLEDGMENTS

We would like to thank Stephen Bamford for his valuable comments on earlier versions of this manuscript and our Data Transparency peers for their input and sharing their insights about the PMDA process.

ABOUT THE AUTHORS

Jean-Marc Ferran leads the PhUSE Data Transparency Working Group since 2014 and represented PhUSE at the EMA Stakeholders Group meetings that took place during the development of Policy 0070 External Guidance. Jean-Marc joined later on the EMA Technical Anonymisation Group in 2017 and was a member of the Health Canada Stakeholders Group for Public Release of Clinical Data. Jean-Marc Ferran is also an Independent Consultant and supports d-Wise Technology as an SME with the development of their Data Anonymization products.

Liz is Senior Director and Global Public Policy Lead at UCB. She is responsible for establishing the strategic framework, guiding principles, and corporate policies that inform clinical trial transparency and responsible data sharing. Liz led the TransCelerate Clinical Data Transparency Data Anonymization sub-team and co-leads the PhUSE Data Transparency workstream. Prior to her role in Data Transparency, Liz had more than 20 years' experience working as a biostatistician in the pharmaceutical industry and has an MSc in Applied Statistics.

The opinions in this paper are our own.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Jean-Marc Ferran
Qualiance ApS
Email: JMF@qualiance.dk
Web: http://www.qualiance.dk

Liz Roberts
Email: liz.roberts@ucb.com