



EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Guidance on the anonymisation of clinical reports for the purpose of publication in accordance with policy 0070

Stakeholder webinar
24 June 2015, London



Presented by Monica Dias
Policy Officer

An agency of the European Union





Scope and objectives

- The Agency developed guidance to pharmaceutical industry on anonymisation of clinical reports, in the context of phase 1 of the policy;
- The guidance aims at assisting companies by recommending methods, techniques and a process that could be applied to clinical reports, for the purpose of achieving adequate anonymisation while retaining a maximum of scientifically useful information on medicinal products for the benefit of the public.





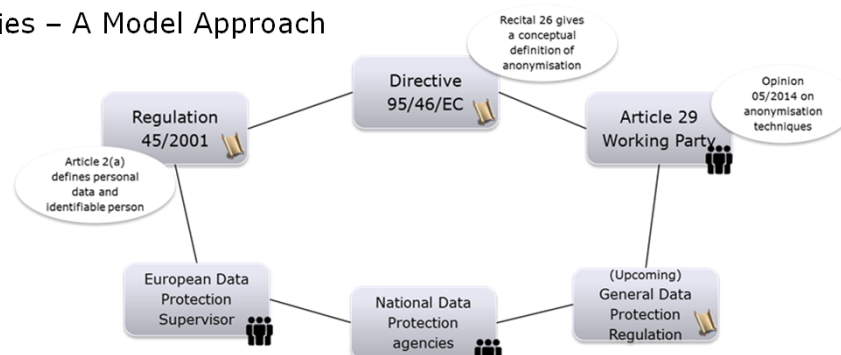
Background

- MAHs/Applicants have the responsibility for submitting clinical reports that have been rendered anonymous for the purpose and use of the clinical reports, i.e. publication under policy 0070 subject to Terms of Use (ToU).
- The anonymised clinical reports should be a copy of the clinical reports submitted in the context of the scientific evaluation procedure, stripped of sufficient elements such that the participants can no longer be identified. The data in the clinical reports must be processed in such a way that it can no longer be used to identify a natural person by using “all the means likely reasonably to be used” by either the controller or a third party, as described in Directive 95/46/EC.



Legal framework and available standards

- EU data protection legislation
- Article 29 Data Protection Working Party opinion of anonymisation techniques (Opinion 05/2014)
- Information Commissioner's Office (ICO) Code of Practice. Anonymisation: managing data protection risk
- Sharing clinical trial data: Maximizing benefits, minimizing risk. Institute of Medicine (IOM)
- Pharmaceutical Users Software Exchange (PhUSE) de-identification standards for CDISC SDTM 3.2
- Transcelerate BioPharma Inc., Clinical Study Reports Approach to Protection of Personal Data and Data De-identification and Anonymisation of Individual Patient Data in Clinical Studies – A Model Approach





Article 29 Working Party Opinion on anonymisation techniques

- Article 29 Opinion on anonymisation provides **two options** to establish if a dataset is anonymised:
 1. Demonstrate that after anonymisation it is no longer possible to:
 - *Singling out*: possibility to isolate some records of an individual in the dataset*;
 - *Linkability*: ability to link, at least, two records concerning the same data subject or a group of data subjects (in the same database or in two different databases);
 - *Inference*: the possibility to deduce, with significant probability, the value of an attribute from the values of a set of other attributes

OR

2. Perform an analysis of re-identification risk.

* In the context of phase 1 of policy 0070, dataset are the set of clinical reports published by the Agency



4 Guidance on the anonymisation of clinical reports for the purpose of publication in accordance with policy 0070



Anonymisation techniques

- Several anonymisation techniques* are available to MAHs/Applicants
- It is a field of active research and rapidly evolving
- From the anonymisation techniques described by the Article 29 Working Party, examples of techniques that could be applicable to clinical reports are:
 - Masking
 - Randomisation - noise addition and permutation (data utility limitations)
 - Generalisation - aggregation and k-anonymity

*The legislation is not prescriptive about the techniques to be used by data controllers.





Anonymisation of *direct* and *quasi identifiers*

- Anonymisation of *direct identifiers*
 - Always removed, e.g. patient ID
- Anonymisation of *quasi identifiers* (not always necessary to redact all quasi identifiers)
 - Dates – individual patient dates can be offset
 - Geographical location - aggregate or generalise from country to region or continent
 - Small populations and rare diseases – risk assessment is key to ensure adequate anonymisation



EMA recommendation to MAHs/Applicants on how to best achieve anonymisation (1/2)

- Guidance is not intended to mandate any specific methodology but to highlight to MAHs/Applicants the available techniques and those the EMA considers most relevant in the context of the anonymisation of clinical reports
 - Masking is likely to be used by MAHs/Applicants initially since pharmaceutical companies will have to anonymise their data retrospectively, i.e. after the clinical report has already been written. However, redaction is more likely to decrease the clinical utility of the data compared to other techniques
 - Therefore, randomisation and generalisation techniques are recommended in order to optimise the clinical usefulness of the information published



EMA recommendation to MAHs/Applicants on how to best achieve anonymisation (2/2)

- It is up to a company taking due account of the ultimate purpose and use of the clinical reports (i.e. publication subject to ToU) on the basis of the guidance made available to decide
 - Which option to use (demonstrate that after anonymisation all three criteria are fulfilled - singling out, linkability and inference, or perform a risk assessment)
 - Which anonymisation techniques to use in order to achieve adequate anonymisation while retaining a maximum of scientifically useful information



Anonymisation process

In order to facilitate the company's approach it is recommended to follow the anonymisation process described below:

1. *Determination of direct identifiers and quasi-identifiers in the dataset*
2. *Identification of possible adversaries and plausible attacks on the data*
3. *Data utility considerations*
4. *Determining the risk of re-identification threshold*
5. *Evaluate the actual risk of re-identification*
6. *Anonymisation methodology*
7. *Documenting the anonymisation methodology and process*



Reporting on the anonymisation process

- **Anonymisation report** to be submitted to the Agency together with the anonymised clinical reports
 - The report should contain information on:
 - The anonymisation process, including the methodology (techniques used and the rationale for using them),
 - Outcome of the analysis of the risk of re-identification or alternatively confirmation that the three criteria for anonymisation have been fulfilled
 - Report to be **published** by the Agency together with the anonymised clinical reports



Redaction of personal data of individuals other than trial participants

- In the context of the implementation of the CT Regulation, the Agency is currently performing a Privacy Impact Assessment (PIA) to establish the functionalities of the database, in particular with regard to data fields to be made publicly accessible.
- The outcome of the PIA will be included in the guidance. However, it is anticipated that personal data of the (principal) investigators and their sites will be made public.