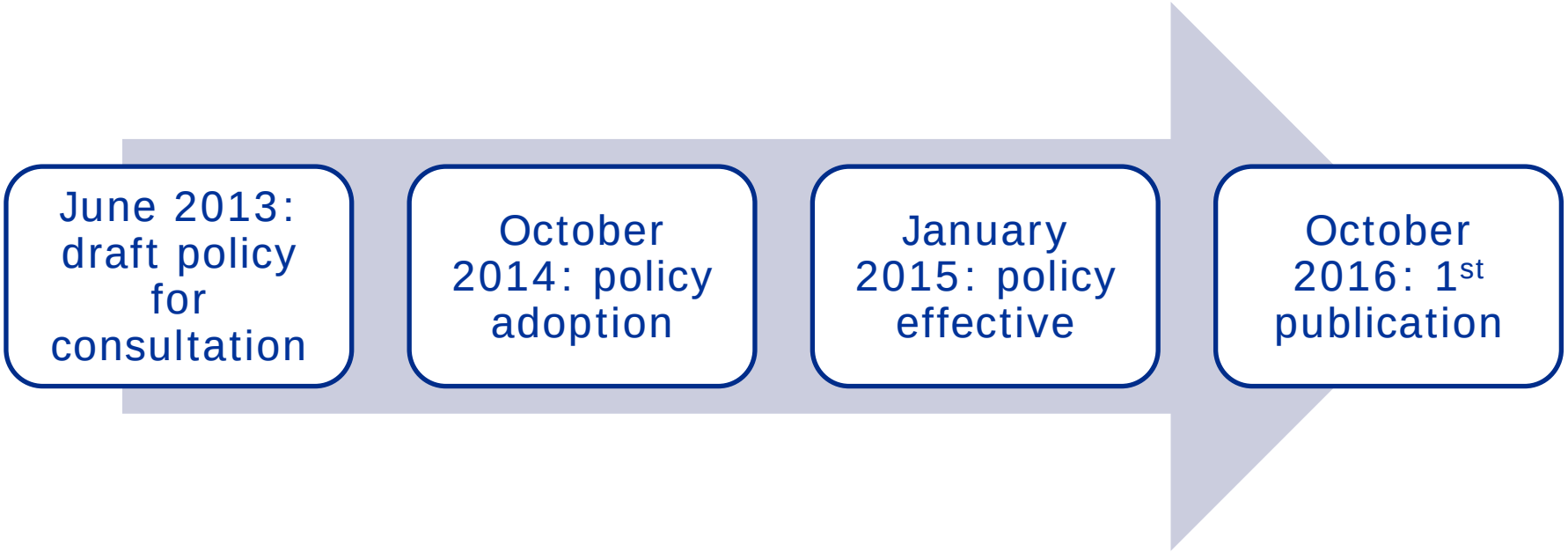
EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Policy 0070: Guidance on Anonymisation

Technical Anonymisation Group (TAG) meeting, London, 29-30 November 2017
Agenda point 05

Presented by Karen Quigley
Clinical Data Publication Manager

Policy 0070 / Clinical Data Publication (CDP)



Policy implementation

Phase I

- Clinical reports = clinical overview, clinical summary, clinical study reports, protocol & amendments, sample case report form, documentation of statistical methods
- **EMA is working on Phase I implementation**

Phase II

- Individual patient data (IPD)
- **Later stage**

Clinical Data Publication Guidance

Introduction, scope, definitions

External guidance on the **procedural aspects** related to the submission of clinical reports for the purpose of publication in accordance with EMA policy 0070

Guidance on the identification and redaction of **commercially confidential information** (CCI) in clinical reports submitted to the EMA

Guidance to pharmaceutical industry on the **anonymisation** of clinical reports for the purpose of publication in accordance with EMA policy 0070

Published on EMA website-rev 3:
http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2017/09/WC500235371.pdf

Guidance

- The Agency published in March 2016, guidance on the anonymisation of clinical reports which provides information to the pharmaceutical industry on the anonymisation of clinical reports (phase 1 - no IPD), i.e. on some of the anonymisation techniques that are available to and can be used in CSR documents.
- The guidance has been updated following webinars with stakeholders in 2016-7.
- The field of anonymisation, and in particular the techniques used by controllers of personal data to anonymise data, is a field of active research and rapidly evolving and the EMA guidance needs to reflect this.

Chapter 3 – guidance on anonymisation of clinical reports

1. Introduction and **2. Legal framework and available standards**

3. General considerations Context of data disclosure - Concept of anonymisation - Anonymisation criteria - Anonymisation techniques - Advances in technology

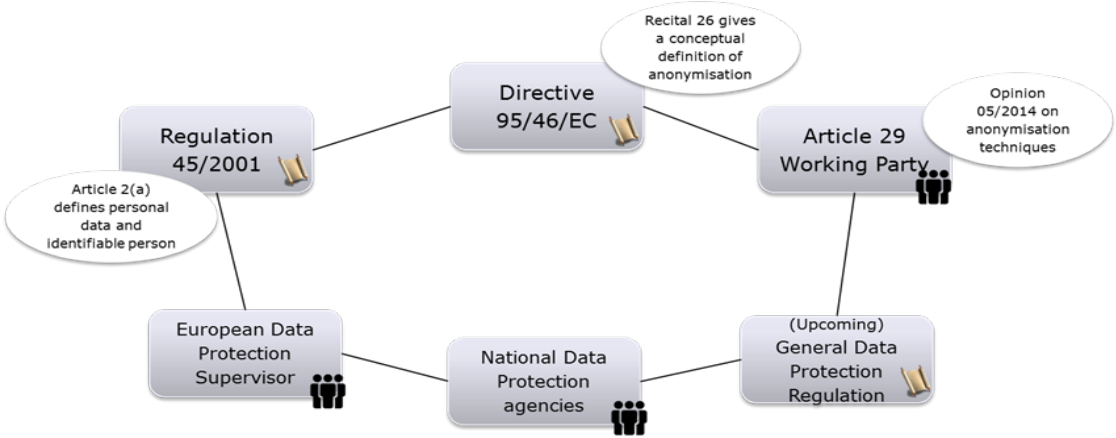
4. Applying these general considerations in the context of clinical reports for publication in accordance with EMA policy Context of data disclosure - Concept of anonymisation - Data utility - Advances in technology - Rare diseases and small populations

5. EMA recommendations to MAHs / applicants on how best to achieve anonymisation Data utility - Rare diseases, small populations and low frequency events - Specific recommendations - Anonymisation process

6. Redaction of personal data of investigators, sponsor staff and applicant / MAH staff

Guidance - Introduction and legal framework

- EU data protection legislation (Directive 95/46/EC); now repealed by the GDPR
- Article 29 Data Protection Working Party opinion of anonymisation techniques (Opinion 05/2014)



Guidance - Introduction and legal framework

- **EMA consulted the EDPS** on Policy 0070, and in particular the **draft External Guidance on Anonymisation**.
- EDPS recognised the inherent conflict between transparency and protection of privacy and praised EMA "*for the considerable effort to provide practical guidance on how to achieve these conflicting objectives*".
- **EDPS acknowledges** that for Policy 0070 there are constraints which make the **very low likelihood of possible re-identification** of the data subject the only consistent safeguards for the protection of personal data.

Guidance – General Considerations

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. Re-identification techniques are evolving and the guidance needs to reflect this and the range of different procedures within scope of the policy.

Two options to establish if data is anonymised

- three criteria from Art 29 WP (singling out-linking-inferring) or
- risk assessment –qualitative or quantitative

Anonymisation techniques

- Masking (removal of values for variables) is likely to be used by MAHs/Applicants initially since companies will have to anonymise their data retrospectively. However, redaction used alone 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.

Anonymisation process

The recommended approach to follow for the anonymisation process is described below:

1. *Documenting the anonymisation methodology and process*
2. *Determination of direct identifiers and quasi-identifiers*
3. *Assessment of anonymisation*
4. *Identification of possible adversaries and plausible attacks on the data*
5. *Data utility considerations*
6. *Evaluation that the actual risk of re-identification is below threshold set after data has been anonymised*

Methods and outcome provided in AnR submitted with clinical documents and is published on clinical data portal

Data utility

- Should be integrated in the risk assessment
- Different techniques lead to different levels of data utility
- Encourage techniques that favour optimising data utility while ensuring data anonymisation is achieved
- Rare diseases, small populations are high risk and anonymisation adapted to the identified risk
- Goal is to balance high utility and low risk re-identification

Guidance – redaction of personal data-investigators sponsors

- Personal data of sponsors/MAH staff and investigators will not be published
- Exception to this - signatories of clinical study report and principal investigator (PI) but contact details and signatures are protected
- PI redacted where geographical information may be linked to patients and lead to re-identification of a patient

Guidance - Issues noted to date

1) What level of technical/statistical information is needed in the AnR so EMA can understand the anonymisation method proposed?

2) On which basis/criteria could one decide whether a data variable (e.g gender) is a quasi-identifier and should be listed in the report?

Possible criteria could be for example: the contribution of the data variable to the risk of re-identification **in general** or the level of knowability (universally known as opposed to known by a few people) of the data variable **in each particular case**.

Guidance- Issues noted to date

3) Analytical or non-analytical approach- terminology to be used

When a qualitative approach is used (3 criteria), we're looking at a non-analytical approach since the company evaluates the risk of re-identification in a crude way and not using analytics (analysing data). Other qualitative approach – risk assessment?

When a quantitative approach is used, we're talking about an analytical approach since the company evaluates the risk of re-identification using analytics (analysing data).

Any questions?

Acknowledge EMA colleagues

European Medicines Agency
30 Churchill Place • Canary Wharf • London E14 5EU • United Kingdom
Telephone +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5555
Send a question via our website www.ema.europa.eu/contact

Follow us on **@EMA_News**