Clinical Study Reports Approach to Protection of Personal Data

Background

TransCelerate BioPharma Inc. is a non-profit organization of biopharmaceutical companies focused on advancing innovation in research and development (R&D), identifying and solving common R&D challenges, thus increasing the quality of clinical studies and delivering more high-quality medicines to patients. The biopharmaceutical members of TransCelerate are committed to enhancing public health and medical and scientific knowledge through the sharing and transparency of clinical trial information.

Introduction

In July 2013 PhRMA and EFPIA member companies committed to "Principles for Responsible Clinical Trial Data Sharing"[1] This included a commitment to enhance public access to clinical study information and clinical study reports (CSRs) as follows:

- For any submissions filed as of January 1, 2014[2], following approval in the US and EU, biopharmaceutical companies will make publicly available, at a minimum the synopses of clinical study reports (CSRs) for clinical trials submitted to regulatory authorities in the US, EU, and to national competent authorities of EU Member States.
- Companies will evaluate requests for full CSRs by researchers / investigators as they do for requests for access to patient-level data, study-level data and protocols for legitimate uses.

These data sharing principles also have applicability in major markets other than the US and EU at the discretion of individual companies.

While companies' approaches to providing CSRs align with these principles, individual companies may differ in: the scope of CSRs provided; when reports are provided; and other aspects. For example some companies will publicly disclose full CSRs while others will publicly disclose CSR synopses and provide full CSRs to other researchers according to specific criteria. Companies will make this information available, but will comply with the need to protect privacy of individuals, groups and staff associated with a clinical study.

[1] PhRMA, EFPIA Principles for responsible clinical trial data sharing. July 18, 2013 http://www.phrma.org/sites/default/files/pdf/PhRMAPrinciplesForResponsibleClinicalTrialDataSharing.pdf

[2] EFPIA Press Release.: http://efpia.eu/mediaroom/132/43/Joint-EFPIA-PhRMA-Principles-for-Responsible-Clinical-Trial-Data-Sharing-Become-Effective

The protection of Commercial Confidential Information (CCI) is a matter for individual companies, and further discussions and considerations are needed on this topic. In contrast, privacy considerations are not company specific and, because the global privacy landscape is diverse, they can be region or country specific. TransCelerate establishes a general approach that can be applied across most CSRs globally, however, adjustments need to be made for local national privacy laws and regulations.

Privacy Considerations and Scope

The TransCelerate members have agreed that privacy concerns are paramount in the context of public disclosure of CSRs. Due consideration has been given to protect the privacy of individuals and groups associated with a clinical study: patients / subjects (also referred to as research participants); investigators, site staff, research institutions and staff; vendor/co-development partner companies and their staff; and sponsor company staff.

Accordingly, the following privacy considerations are being applied to the sharing of CSRs:

- The standard approach seeks to protect privacy regardless of the audience.
- The extent to which the content of a CSR is disclosed is at the discretion of an individual company and may differ for current and legacy CSRs. If content is disclosed, then the agreed standards apply to all sections disclosed. This is an evolving approach.
- Case-by-case assessments may be required by individual companies to determine the appropriateness of disclosing a particular CSR in special circumstances (e.g., rare diseases, small populations), where any disclosure may jeopardize privacy. Additionally, for individual patient cases where normal redaction approaches would have a higher risk of re-identification, these would also require case-by-case assessment (e.g. survival data in a small subset of patients, a single subject with complete response).
- Aggregated data or descriptions of aggregated data and study level information (e.g. public register IDs, tabular, graphic, or cross-patient data) will rarely raise privacy concerns and can be retained. An exception might be aggregate categories with a single subject and a reasonable likelihood of being re-identified .

Anonymized datasets and CCI are out of scope for the standard approach described in this consensus document.

General Approach

In keeping with the principles of protecting privacy, all personally identifiable information (PII) within a CSR is de-identified by removal (i.e., certain sections of the CSR are removed) or redaction (i.e., specific content is masked irreversibly from view, for example with a black bar) prior to CSR disclosure. An attempt should be made to explain why information has been removed. Appropriate explanatory text should be provided in a manner and location selected by the sponsor (see Appendix 1 for examples).

The assessment of what information to remove or redact has been based upon the determination of the potential risk of jeopardizing personal privacy by disclosing the information. Consideration was also given to direct and indirect personal identifiers suggested by the Health Insurance Portability and Accountability Act (HIPAA[3]) and EU data protection regulations No. 45/2001[4] and Data Protection Directive 95/46/EC[5], which define personal data broadly to encompass any information relating to an identified or identifiable natural person. Also considered were the minimum standard for de-identifying data as described in Hrynaszkiewicz[6]. Sections of CSRs that contain many identifiers, such as investigator curriculum vitae (CVs) / biographies, present a high risk to personal identification and these sections are removed in their entirety. For information within sections of a CSR that present a lower risk to privacy, such as a signature, these are redacted as individual pieces of information.

What is Removed and Redacted

Details of the approach for individual CSR sections and associated information are described below :

- Appendix 2: **Summary of What is Removed and Redacted** - By category: patient / subject; investigator, research institution and their staff; vendor / co-development partner company and staff; sponsor company staff; and study level
- Appendix 3: **Summary of Approach - By clinical study report section** an illustrative implementation aid, which will need to be refined according to individual company CSR practices

[3] HIPAA Privacy Rule
http://privacyruleandresearch.nih.gov/pr_08.asp#8a

[4] Regulation (EC) No. 45/2001 of the European Parliament and of the Council of 18 December 2000 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:008:0001:0022:EN:PDF

[5] Directive 95/46/EC of the European Parliament and of the Council of October 24, 1995 on the protection of individuals with regards to the processing of personal data and on the free movement of such data. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1995L0046:20031120:EN:PDF

[6] Hrynaszkiewicz, I., M. L. Norton, et al. (2010). 'Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers.' BMJ 340: c181.
http://www.bmj.com/content/340/bmj.c181#alternate1

5.1 Removal

- Full patient narratives and corresponding forms (e.g., CIOMS) are **removed**.
- Listings of individual patient data are **removed**.
- Any figures (or tables) that include information pertaining to a single individual are either **removed, or are retained** with the subject identification number **redacted** (e.g. individual subject pharmacokinetic plots)**,** provided this does not jeopardize privacy.

Investigators' CVs / biographies are **removed**.

5.2 Redaction

- **Personal Information**: Names, initials, email addresses, phone and fax numbers, academic/organizational titles, and scanned signatures of all those involved in the study are **redacted** for e.g.
  - patient / subject
  - site staff, research institution and staff. Members of Institutional Review Boards / Independent Ethics Committees (IRB/IECs), and members of committees such as Independent Data Monitoring Committees and Statistical Data Analysis Committees (IDMCs/SDACs).
  - vendor / co-development partner staff
  - company staff (including monitors listed as emergency contacts in the protocol and names included on citations / references for internal company reports).

Any additional personal information (e.g., facial photographs or comparable images) are also **redacted**.

- **Investigator names**: should be **redacted** unless appropriate contractual agreements are in place to retain them.
- **All non-sponsor company names**: research institution, vendor / co-development partner should be **redacted** unless appropriate contractual agreements are in place to retain them or the partnership is publicly known.
- **Addresses**:
  - All patient addresses should be **redacted** in full.
  - All sponsor addresses should be **retained** in full.
  - All other addresses (investigator, research institution, vendor / co-development partner, IRB/IECs, and IDMCs/SDACs) should be **redacted**, retaining **country**.
- **Patient or subject ID numbers** are **redacted.**
- **Individual outcomes**: ID numbers and the associated description of an individual patient's medical outcome are redacted (e.g., outcome of pregnancy, rare disease or event, congenital abnormality) or sensitive data[6] (e.g. illicit drug use or "risky behavior") within the body of the CSR, or in a table footer. Further redaction of the outcomes should be assessed by the individual company based on the risk of re-identification.

- **Verbatim text** in the context of an individual patient is **redacted**; note that coded terms for adverse events, serious adverse events, medical history, and concomitant medications are **retained**. Verbatim text may be **retained** provided a case-by-case assessment for re-identification potential is undertaken.
- **Demographic characteristics**: patient-level data (sex, age, race, ethnicity, height, weight) or socioeconomic information (such as occupation) are **redacted**.
- **All dates relating to an individual patient** are redacted entirely, for example:
  - Dates of birth.
  - Event or assessment dates for individual patients/subjects.
- **Unique identifying numbers** (e.g., Staff User ID, Investigator ID, Laboratory ID, Site/Center ID, Case ID, Randomisation / Treatment Number, Manufacturing Control Numbers) are **redacted**.

What is Retained

The other parts of the CSR are retained including:

- Original Table of Contents and Bookmarks (with the exception of any patient-identifying information).
- Tables or figures containing summary information (with the exception of any patient-identifying information).
- Study roles (e.g., Principal Investigator, Statistician, sponsor responsible medical officer).
- Academic qualifications (e.g., MD, PhD).
- Country location for investigator, non-sponsor vendor / co-development partner companies, IRB/ECs, and IDMCs/SDACs.
- References to the sponsor company including addresses (references to the sponsor would include vendors or CROs who take responsibility for conducting nonclinical or clinical studies).
- Study ID.
- Document Control Numbers (DCNs) linked to any company internal documentation (e.g. HM2007/00444/00) (Individual company discretion to redact if concern that it may jeopardise internal links to systems and understanding that internal links to systems would be broken).
- Public register ID numbers such as IND, NDA, EudraCT, NCT number.
- Published citations and references including names of journals and authors. In citations of internal company reports, names of company authors are **redacted**.
- Dates not related to study participants (e.g. study report dates, signature dates, study milestones including First Subject First Visit, First Blood Sample Collected, Last Subject Last Visit).

Technical Recommendations

For documents existing in an electronic format, we recommend redaction utilizing a professional redaction tool that prevents the ability to revert content and unmask redacted information. Text is irreversibly masked (and rendered invisible for search capabilities) and replaced with a box or text obliterating the information using a redaction tool. For example, a text box or black box:

Personal information included in meta-data should also be removed.

Quality control measures must be employed before the redactions are applied and the CSR is rendered as a Portable Document Format (PDF) that cannot be edited.

For documents not available in an electronic format, alternate redaction methods with the same desired outcomes should be used.

Appendix 1: Example Explanatory Text

At the discretion of individual companies

Category	Standard Text
Key Messages	• Information will be removed or redacted in order to protect the privacy of patients and all named persons associated with the study. • Patient data listings will be completely removed to protect patient privacy. Anonymized data from each patient may be made available subject to an approved research proposal. • Aggregate data will be included; with any direct reference to an individual patient excluded.
Patient Data Listings / Figures and Tables containing data pertaining to an individual patient/subject	This section contained data from each individual patient, rather than in aggregate. They have been excluded to protect patient privacy. Anonymized data from each patient may be made available subject to an approved research proposal.
Full Patient / Case Narratives	This section contained patient narratives which are textual descriptions of medical history, treatment and outcome for individual patients who experienced a clinically important adverse event including serious adverse events during the trial. They have been excluded to protect patient privacy. This data may be made available subject to an approved research proposal and a determination of the ability to provide information from the specific narratives while protecting the patient's privacy.
Investigators' Curriculum Vitae / Biographies	This section contained Principal Investigator's Curriculum Vitae and has been excluded to protect Principal Investigator privacy.

Appendix 2: Summary of What is Removed and Redacted

Category	Remove	Redact	Retain
Patient/Subject	• Patient- level data listings • Full patient narratives and corresponding forms (e.g., CIOMS)	• Patient or subject ID numbers are redacted. • Redact subject ID numbers associated with any description of an individual patient's medical outcome (e.g., outcome of pregnancy, rare disease or event, congenital abnormality) or sensitive data (e.g. illicit drug use or "risky behavior") within the body of the CSR, or in a table footer. Further redaction of the outcomes should be assessed by the individual company based on the risk of re-identification. • Personal information: name, initials, email, phone number, signature, full address including country. • All dates relating to an individual patient are redacted entirely. • Randomization / Treatment number. • Case ID. • Verbatim text. • Patient level demographic information (sex, age, race, ethnicity, height, weight) or socioeconomic information (such as occupation).	• Tables or figures containing summary information (with the exception of patient identifying information)

Recommended Approach: CSR Redaction of Privacy Information

Category	Remove	Redact	Retain
Investigator, Research Institutions and their staff including non-sponsor members of IRB, Ethics Committee and IDMC/SDAC	• Investigators' Curriculum Vitae / Biographies	• Investigator ID and Site/Centre ID • Investigator names, should be redacted unless appropriate contractual agreements are in place to retain them. • Personal information: initials, email, phone number, signature, academic (e.g. Professor) or organizational (e.g. VP of Oncology Research) titles, address excluding country • IRB, Ethics Committee name and IDMC/SDAC: committee and members names, email, phone number signatures, academic (e.g. Professor) or organizational (e.g. VP of Oncology Research) titles , address, excluding country	• Study roles (e.g. Principal Investigator, Statistician) • Academic qualifications (e.g., MD, PhD) • Country for investigator, research institution, IRBs/ECs, and IDMCs/SDACs

Recommended Approach: CSR Redaction of Privacy Information

Category	Remove	Redact	Retain
Vendor (any service provider) or co-development partner company and staff		• Personal information: name, initials, email, phone number, signature, academic (e.g. Professor) or organizational (e.g. VP of Oncology Research) titles, address excluding country • Non-sponsor vendor/co-development partner company names and addresses excluding country, Company names should be redacted unless appropriate contractual agreements are in place or partnership is know publicly.	• Study roles (e.g. Principal Investigator, Statistician) • Academic Qualifications (e.g. MD) • Country for non-sponsor vendor/co-development partner company
Sponsor company staff including sponsor company members of IDMC/SDAC		• Staff User ID • Personal information: name, initials, email, phone number and signature, academic (e.g. Professor) or organizational (e.g. VP of Oncology Research titles)	• Study roles (e.g. Principal Investigator, Statistician, sponsor responsible medical officer) • Academic Qualifications (e.g. MD) • References to the sponsor company • Sponsor company addresses

Recommended Approach: CSR Redaction of Privacy Information

Category	Remove	Redact	Retain
Study Level	• Tables / figures with individual patient level data are removed, or are **retained** with subject ID redacted	• Manufacturing Control Numbers and Laboratory ID • In citations of internal company reports, names of company authors are redacted.	• Original Table of Contents and Bookmarks (with the exception of patient identifying information) • Study ID. • Document Control Numbers (DCNs) linked to any company internal documentation, e.g. HM2007/00444/00. • Public register ID Numbers such as IND, NDA, EudraCT, NCT number • Published citations and references including names of journals and authors. • Dates not related to study participants (e.g. study report dates, signature dates, study milestones including First Subject First Visit, First Blood Sample Collected, Last Subject Last Visit)

Appendix 3: Summary of Approach by Clinical Study Report Section

This table is an illustrative implementation aid, which will need to be refined according to individual company CSR practices. (Table in excel format available upon request)

Content of Clinical Study Report	Approach													
	Remove				Redact									
	5.1.1 Patient narratives	5.1.2 Listings of patient level data	5.1.3. Figures with patient level data	5.1.4. Investigator CV's and biographies	5.2.1 Personal information	5.2.2 Investigator names*	5.2.3 Non-sponsor company names*	5.2.4 Addresses	5.2.5 Patient or subject ID #s	5.2.6 Individual subject outcomes	5.2.7 Patient level verbatim text	5.2.8 Patient level demographic characteristics	5.2.9 Patient level dates	5.2.10 Unique IDs
Responsible medical officer signature page Report author page					√									
Investigator signature page					√	√		√						
Cover / title page					√		√	√						
Table of Contents (and Bookmarks)					√	√	√		√					
Abbreviations Glossary of terms Trademarks														
Ethics and Good Clinical Practice														
Synopsis (except publically available)					√		√	√	√		√	√	√	
Introduction														
Investigators and Study Administrative Structure					√	√	√	√						√
Investigational plan							√	√						
Results, Including: • Study population • Efficacy • Safety (Including serious adverse events, other significant adverse events, pregnancy) • Immunogenicity • Pharmacokinetic • Pharmacodynamic and biomarketer • Relationship between pharmacokinetic and pharmacodynamic parameters									√	√	√	√	√	
Discussion and Conclusions									√	√	√	√	√	
References					√									
Internal report														
Publication														
Post-text tables, figures, and listings		√	√						√	√	√	√	√	√
Case Narratives	√													
Protocol and Protocol Amendments					√	√	√	√						
Statistical Analysis Plan					√		√	√						
Sample Case Report Form, Diary Cards and Questionnaires					√	√	√	√						
List of IECs or IRBs, List of Principal and sub-investigators, other important participants in the study					√	√	√	√						√
ICH Specified Data Listings		√							√					√
Other Data Listings		√							√					√
Publications based on the study														
Sample Consent Form and Written Information for Patients					√	√	√	√						
Principal Investigators' Curriculum Vitae/Biography				√										
Listing of subjects receiving investigational product(s) from specific batches, if more than one batch was used.									√					√
Randomization scheme and codes									√					√
Audit Certificates (Note: these are internal company audits)					√	√	√	√						
Documentation of Interlaboratory Standardization Methods and Quality Assurance Procedures (if used)					√	√	√	√						
Independent Data Monitoring Committee Information					√	√	√	√						
Statistical Methods					√	√	√	√						
Bioanalytical Methods and Results, Including: • Pharmacokinetic or Population Pharmacokinetic Report • Biomarker Report • Pharmacogenetic Results • Health Outcomes Report on Direct Cost Data • Virology Genotypic and Phenotypic Results	√	√	√	√	√	√	√	√	√	√	√	√	√	√

Recommended Approach: CSR Redaction of Privacy Information

Summary of Approach by Clinical Study Report Section Table Continued

Content of Clinical Study Report		Retain										
	6.0 No Privacy information	6.1 Original ToC and bookmarks	6.2 Summary tables and figures	6.3 Study roles	6.4 Academic qualifications	6.5 Country locations	6.6 References to the sponsor	6.7 Study ID	6.8 DCNs	6.9 Public register IDNs	6.10 Published citations	6.11 Dates not related to study participants
Responsible medical officer signature page Report author page				√	√							√
Investigator signature page				√	√	√						√
Cover / title page				√	√	√	√					√
Table of Contents (and Bookmarks)		√										
Abbreviations Glossary of terms Trademarks	√											
Ethics and Good Clinical Practice	√											
Synopsis (except publically available)				√	√	√	√					√
Introduction	√											
Investigators and Study Administrative Structure				√	√	√						√
Investigational plan						√						
Results, Including: • Study population • Efficacy • Safety (Including serious adverse events, other significant adverse events, pregnancy) • Immunogenicity • Pharmacokinetic • Pharmacodynamic and biomarketer • Relationship between pharmacokinetic and pharmacodynamic parameters												
Discussion and Conclusions												
References											√	
Internal report												
Publication												
Post-text tables, figures, and listings			√									
Case Narratives												
Protocol and Protocol Amendments				√	√	√	√					√
Statistical Analysis Plan				√	√	√	√					√
Sample Case Report Form, Diary Cards and Questionnaires				√	√	√	√					√
List of IECs or IRBs, List of Principal and sub-investigators, other important participants in the study				√	√	√						√
ICH Specified Data Listings			√									
Other Data Listings			√									
Publications based on the study												
Sample Consent Form and Written Information for Patients				√	√	√	√					√
Principal Investigators' Curriculum Vitae/Biography												
Listing of subjects receiving investigational product(s) from specific batches, if more than one batch was used.												
Randomization scheme and codes												
Audit Certificates (Note: these are internal company audits)				√	√	√	√					√
Documentation of Interlaboratory Standardization Methods and Quality Assurance Procedures (if used)				√	√	√	√					√
Independent Data Monitoring Committee Information				√	√	√	√					√
Statistical Methods				√	√	√	√					√
Bioanalytical Methods and Results, Including: • Pharmacokinetic or Population Pharmacokinetic Report • Biomarker Report • Pharmacogenetic Results • Health Outcomes Report on Direct Cost Data • Virology Genotypic and Phenotypic Results		√	√	√	√	√	√	√	√	√	√	√