

Data management in health technologies

Keynote

Sven Trelle, CTU Bern

02.03.2021

u^b

^b
UNIVERSITÄT
BERN



**BIG DATA
CAN GENERATE
BIG BRAINSTORMS**

BIG DATA IS JUST A FAD



**IT'S JUST
1960's BATCH
PROCESSING,
BUT IN COLOR**

Topics to be covered

- Regulatory framework
- Other requirements and standards
- Some aspects of good science in the 21st century ...



- All content based on **CTU-interpretation** of the Human Research Act, its ordinances, and related documents (e.g. Botschaft, Erläuterungen etc.)
- Interpretation of the act and its ordinances not 100% harmonized across different authorities
- Although all content was checked carefully: Errors and Omissions Excepted (E&OE)



Scope

Art. 2 Scope

¹ This Act **applies** to research concerning **human diseases** and concerning the **structure** and **function** of the **human body**, which involves:

- a. **persons**;
- b. deceased persons;
- c. embryos and foetuses;
- d. **biological material**;
- e. health-related personal **data**.

² It **does not apply** to research which involves:

- a. IVF embryos in accordance with the Stem Cell Research Act of 19 December 2003¹;
- b. **anonymised** biological material;
- c. **anonymously collected** or anonymised health-related data.

Human Research Act

Development over time

2010

- Constitution

2011

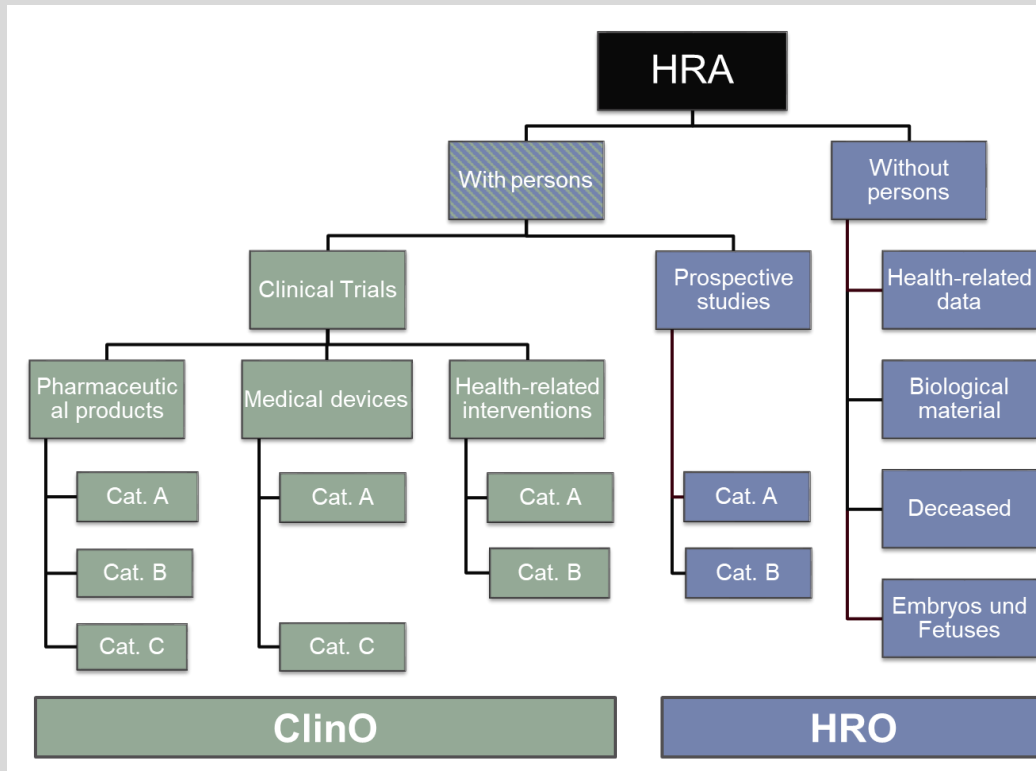
- Human Research Act (HRA)
Explanation in «*Botschaft*»

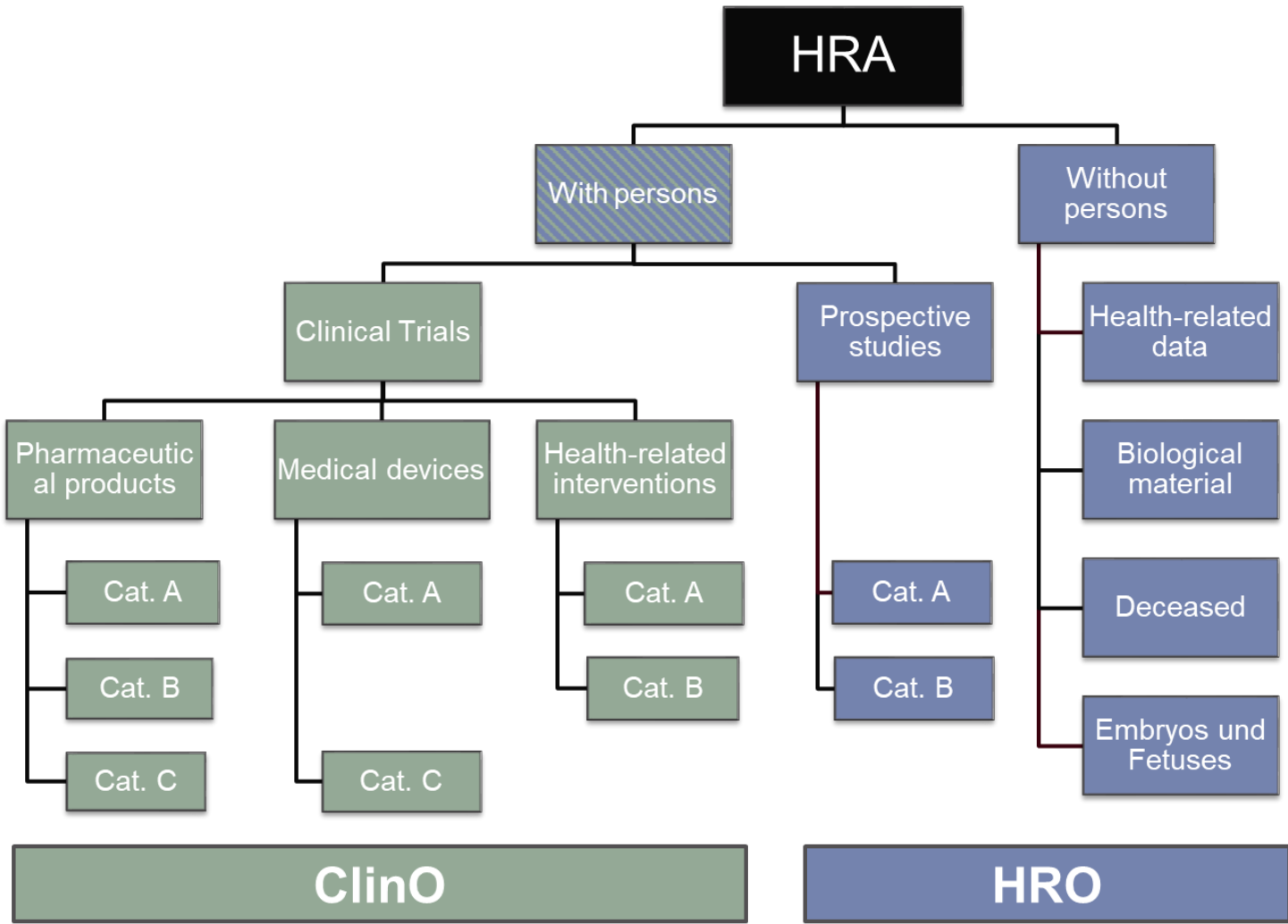
2014

- ClinO
 - HRO
 - OV-HFG
- } • Explanation in «*Erläuterungen*»
• *Rütsche B (Hrsg). Humanforschungs-gesetz (HFG). Bern: Stämpfli. 2015.*

Human Research Act

Overview





Federal Act on Data Protection (FADP)

Sensitive personal data

Art. 3 Definitions

The following definitions apply:

...

c. *sensitive personal data*: data on:

1. ...
2. health, the intimate sphere or the racial origin,
3. social security measures,
4. ...

...

Data storage

Art. 5 Storage of health-related personal data and biological material

¹ Any person who stores health-related personal data for research must take appropriate operational and organisational measures to protect it, and in particular:

- a. restrict the handling of the health-related personal data to those persons who require this data to fulfil their duties;
- b. prevent unauthorised or accidental disclosure, alteration, deletion and copying of the health-related personal data;
- c. document all processing operations which are essential to ensure traceability.

² ...

Proportionality principle

HRO Art. 5, Para. 2, Let. a

- HRO in-line with data protection regulations e.g. Datenschutzgesetz Bern (KDSG) Art. 5

³ Die Personendaten und die Art des Bearbeitens müssen für die Aufgabenerfüllung **geeignet und notwendig sein.**

- Concerns mainly access to identifying information (names, contact details, AHV No. etc.)
- But applies also to the data as such

Laissez-faire?

Practice



The screenshot shows the website of the Swiss Federal Assembly (Die Bundesversammlung — Das Schweizer Parlament). The navigation bar includes links for 'LEICHTE SPRACHE', 'PARLNET', 'KONTAKT', and language options 'DE', 'FR', 'IT', 'RM', 'EN'. Below the navigation bar are categories like 'ORGANE', 'RATSBETRIEB', 'ÜBER DAS PARLAMENT', 'SERVICES', and 'INTERNATIONALES'. A search icon is visible on the right. The breadcrumb trail reads: '... > SERVICES > NEWS > DATENSCHUTZGESETZ: KEIN KOMPROM...'. The main content area features the date 'Freitag, 03. Juli 2020 11h00' and the title 'MEDIENMITTEILUNG' followed by the main headline 'DATENSCHUTZGESETZ: KEIN KOMPROMISS BEIM PROFILING'. Below the headline, the text begins: 'Eine der noch nicht beseitigten Differenzen zwischen Nationalrat und Ständerat bei der Revision des Datenschutzgesetzes betrifft das Profiling. Die Staatspolitische Kommission des Nationalrates lehnt die vom Ständerat verabschiedete Kompromisslösung ab und beantragt ih-'. On the right side, there is an 'AUTOR' section with a person icon, listing 'SPK-N Sekretariat der Staatspolitischen Kommissionen CH-3003 Bern'.

Die Bundesversammlung — Das Schweizer Parlament

LEICHTE SPRACHE PARLNET KONTAKT DE FR IT RM EN

ORGANE RATSBETRIEB ÜBER DAS PARLAMENT SERVICES INTERNATIONALES

... > SERVICES > NEWS > DATENSCHUTZGESETZ: KEIN KOMPROM...

Freitag, 03. Juli 2020 11h00

MEDIENMITTEILUNG

DATENSCHUTZGESETZ: KEIN KOMPROMISS BEIM PROFILING

Eine der noch nicht beseitigten Differenzen zwischen Nationalrat und Ständerat bei der Revision des Datenschutzgesetzes betrifft das Profiling. Die Staatspolitische Kommission des Nationalrates lehnt die vom Ständerat verabschiedete Kompromisslösung ab und beantragt ih-

AUTOR 

SPK-N
Sekretariat der Staatspolitischen
Kommissionen
CH-3003 Bern

Views on data protection/privacy

Personal experience only!

- Clinician-researchers are relatively liberal (naïve) when it comes to data protection and clinical research
- Informal proof: CTU Bern regularly receives datasets with directly identifying information e.g. names, date of birth (at least monthly) and there is lack of understanding on the other side if we reject these datasets and ask them to de-identify

HRO Art. 5, Para. 2, Let. b

b. prevent unauthorised or accidental disclosure, alteration, deletion and copying of the health-related personal data

- Location of storage: central storage/server
- Access: user management
- Internal standards and defined processes: quality management and standard operation procedures

- Prospective studies (data entry): dedicated study management system with a data model
 - Frontend i.e. electronic case report forms
 - Backend i.e. relational database

HRO Art. 5, Para. 2, Let. c

c. document all processing operations which are essential to ensure traceability.

- Internal standards and defined processes: quality management and standard operation procedures
- Audit trail or version control: who, when, what

Simply administrative burdens?

Complaints everywhere

FMH Aktuell

4

Repräsentative Befragung der Ärzteschaft im Auftrag der FMH

Der administrative Aufwand der Ärzterschaft nimmt weiter zu

Bruno Trezzini^a, Beatrix Meyer^b, Melanie Ivankovic^c, Cloé Jans^d, Lukas Golder^e

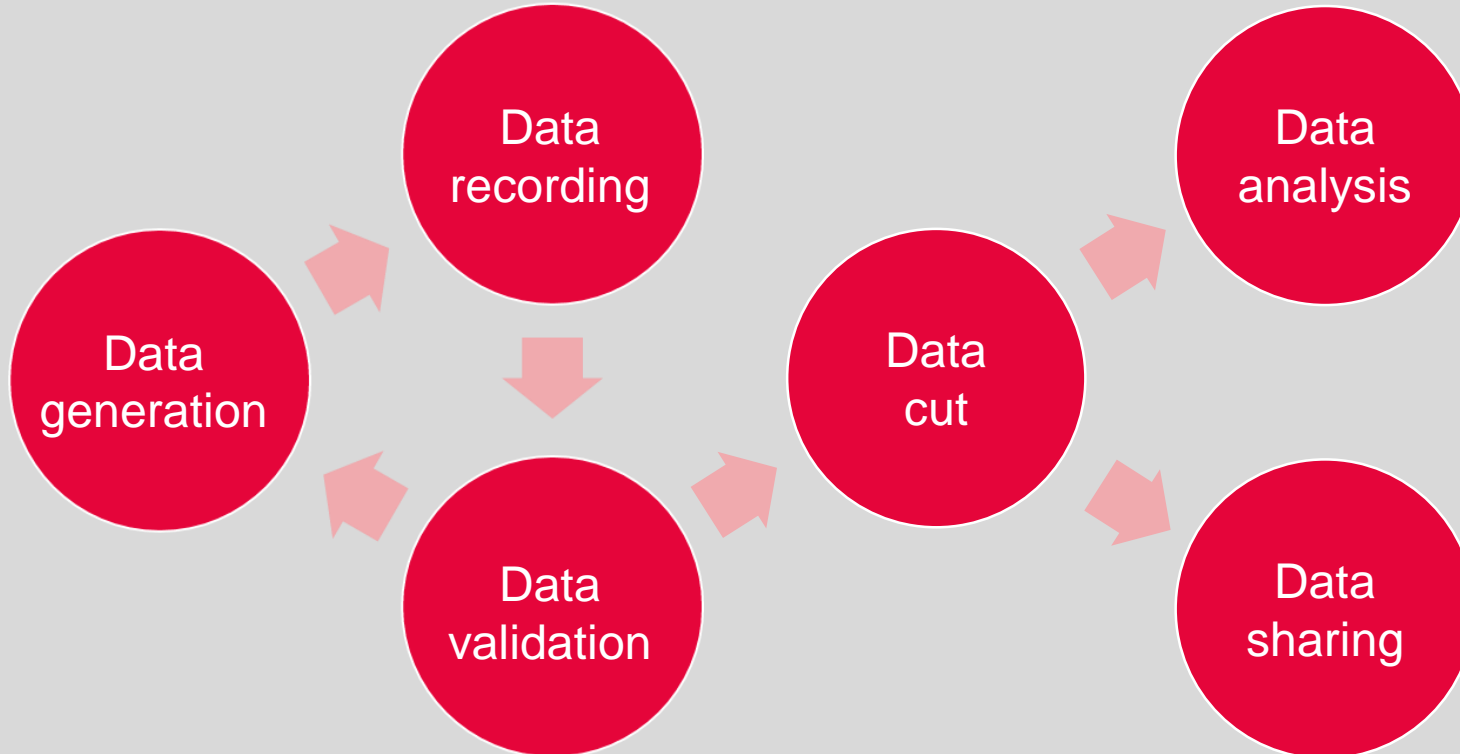
^a Dr. phil., wissenschaftlicher Mitarbeiter, Abteilung Stationäre Versorgung und Tarife, FMH; ^b Leiterin Abteilung Stationäre Versorgung und Tarife, FMH;

^c Junior Projektleiterin gfs.bern; ^d Leiterin operatives Geschäft gfs.bern; ^e Co-Leiter gfs.bern

Der administrative Dokumentationsaufwand hat sich für die Spitalärzteschaft

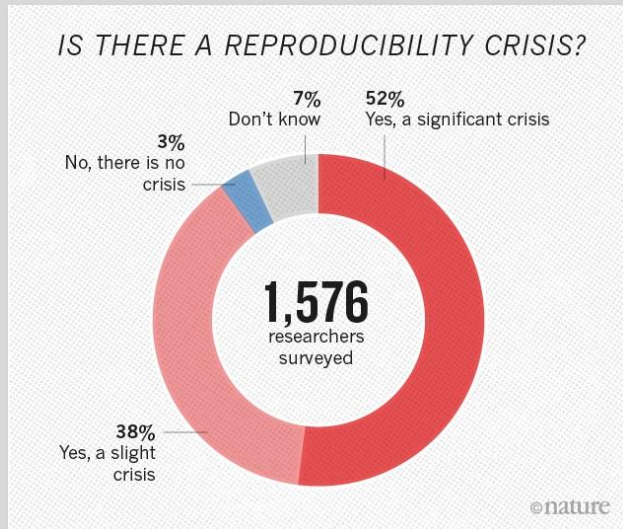
Traceability

In practice



Reproducibility etc.

Scientific research



Recent changes
Upload file
Tools
What links here
Related changes
Special pages
Permanent link
Page information

Article [Talk](#)

Not logged in [Talk](#) [Con](#)

[Read](#) [Edit](#) [View history](#)

Replication crisis

From Wikipedia, the free encyclopedia

The **replication crisis** (also called the **replicability crisis** and the **reproducibility crisis**) is an ongoing [methodological crisis](#) in which it has been found that many scientific studies are difficult or impossible to [replicate](#) or [reproduce](#). The replication crisis most severely affects the [social sciences](#) and [medicine](#).^{[2][3]} The phrase was coined in the early 2010s^[4] as part of a growing awareness of the problem. The replication crisis represents an important body of research in the field of [metascience](#).^[5]

Because the reproducibility of experimental results is an essential part of the [scientific method](#),^[6] an inability to replicate the studies of others has potentially grave consequences for many fields of science in which significant theories are grounded on unreproducible experimental work. The replication crisis has been particularly widely discussed in the fields of medicine, where a number of efforts have been made to re-investigate classic results, to determine both the reliability of the results and, if found to be unreliable, the reasons for the failure of replication.^{[7][8]}

Contents [\[hide\]](#)

- Scope
 - Overall
 - In psychology
 - Psychology replication rates
 - A disciplinary social dilemma
 - "Methodological terrorism" controversy
 - In medicine
 - In marketing
 - In economics

Ioannidis
Publishes
False"^[1]

Baker M 2006 and https://en.wikipedia.org/wiki/Replication_crisis

Terminology

- Reproducibility
- Replicability
- Robustness (sensitivity)
- Generalizability

		Data	
		Same	Different
Analysis (code)	Same	Reproducible	Replicable
	Different	Robust	Generalisable

Kirstie Whitaker (<https://youtu.be/NDNYPDm1-2c>)

Simple trials?

What do we mean by simple, though?

WHY DO WE NEED SOME LARGE, SIMPLE RANDOMIZED TRIALS?

SALIM YUSUF* RORY COLLINS AND RICHARD PETO

Clinical Trial Service Unit, Radcliffe Infirmary, Oxford, UK

The criteria for a good trial are similar in many serious diseases: first and foremost, ask an 'important' question and, secondly, answer it 'reliably'. These two very general criteria obviously require further elaboration, but even as they stand they can suggest some surprisingly specific consequences for clinical trial design. Particularly, they can be used to suggest both the possibility and the desirability of *large, simple randomized* trials of the effects on *mortality* of various *widely practicable* treatments for *common* conditions.

There are six main steps in the argument leading to this conclusion. First, the identification of

How clinical research is organized

The problem with trust and good intentions



Trials versus (?) observational

- Dedicated personnel at sites
- Embedded in routine practice
- Regular monitoring of the data
- Feedback and contact with sites
- Reporting

GARFIELD-AF register

Aims: *The value of data obtained depend critically upon robust quality standards (including source data verification [SDV] and training); features that are commonly absent from registries.*

Conclusion: *The quality standards in GARFIELD-AF have the potential to inform a future 'reference' for registries.*

Documentation and planning

Plans, plans, plans ...

Applications and Projects

Grant application 20

- 1. Personal data
 - Responsible applicant
 - Other applicants
 - Applicants' employment
 - Project partners
- 2. Application data
 - Basic data I
 - Basic data II
 - Link to other SNSF projects
 - Further requested and available funds (not from the SNSF)
 - University or research institution
 - Requested funding
 - Data management plan (DMP)**
 - Research requiring authorisation or notification
 - Exclusion of external reviewers
 - General remarks on the project
- 3. Annexed documents (upload)
 - Research plan
 - CV and publication list
 - Quotes
 - Cover letter
 - Official certificates
 - Other annexes
 - Administrative part of the application

Information/documents

Manage authorisations

Grant application 20 New application
Special Call on Coronaviruses

Start: - In preparation

Data management plan (DMP)

[Integrate existing DMP](#)

Due to the special situation and the short application period, the detailed DMP is only required if the application is accepted. For the submission of the application, a few words for each question are sufficient.

Please describe how you plan to make the research data Findable, Accessible, Interoperable and Reusable (FAIR data principles) in the following sections. Each of the four topics should be addressed with a level of detail appropriate to the project and research field. Sub-questions and help texts are available for each issue. The "questions you might want to consider" will help you to complete the form. However, depending on the project and research field, you may not need to address each of these questions in your DMP.

Complete the DMP form in the same language as your research plan.

The information provided in this template is not part of the scientific evaluation and will not be shared with external reviewers.

Detailed [guidelines](#) are available about the DMP. Furthermore, answers to a set of [frequently asked questions \(FAQs\)](#) about open research data (ORD) are also available.

I do not submit a DMP for the following reason:

1. Data collection and documentation

- 1.1 What data will you collect, observe, generate or reuse?
- 1.2 How will the data be collected, observed or generated?
- 1.3 What documentation and metadata will you provide with the data?

2. Ethics, legal and security issues

- 2.1 How will ethical issues be addressed and handled?
- 2.2 How will data access and security be managed?
- 2.3 How will you handle copyright and Intellectual Property Rights issues?

3. Data storage and preservation

- 3.1 How will your data be stored and backed-up during the research?
- 3.2 What is your data preservation plan?

(Internal) Data Management (Quality) Plan

A (version) controlled document

Template for a Data Management Quality Plan

1. Purpose of Data Management Quality Plan

The purpose of a data management quality plan (DMQP) is to document the life cycle of data collected and the measures to assure the quality of the data throughout its life cycle. It outlines how data will be generated, collected, documented, preserved, and shared during a research project, and after it is completed.

The following topics are specified in the DMQP:

- The nature and context of the research project
- The type of research data that is collected and produced
- The formats, metadata and standards used
- Storage, backups and data security
- Ethical and legal issues
- Access, sharing and reuse of data
- Archiving, storage and long-term preservation
- The roles and responsibilities of the various stakeholders

Table of content

Abbreviations	2
Definitions	2
Dynamic References	2
1. Introduction	2
2. Responsibilities	2
3. Description of collected/generated data	2
4. Case Report Form development.....	2
5. Clinical Data Management System– study specific implementation.....	2
5.1 Implementation of the study database in the Clinical Data Management System	2

5.	Clinical Data Management System– study specific implementation.....	2
5.1	Implementation of the study database in the Clinical Data Management System	2
5.1.1	Codebook development.....	2
5.1.2	Clinical Data Management System implementation	2
5.1.3	Medical coding	2
5.1.4	Data import.....	2
5.2	Verification of Clinical Data Management System setup and deployment.....	2
5.3	Change management	2
6.	Clinical Data Management System – infrastructure	2
6.1	Data storage.....	2
6.2	Data back-up.....	2
6.3	Access to the data	2
6.4	Granting access to the productive version of the Clinical Data Management System and database.....	2
7.	Data collection.....	2

7.	Data collection	2
7.1	Pre-requisites for data entry	2
7.1.1	Data entry guidelines	2
7.1.2	Training of users and training documentation	2
7.2	Entering data	2
8.	Quality control procedures	2
8.1	Real-time data validation	2
8.2	External data validation (offline checks)	2
8.3	Central data monitoring	2
8.3.1	Definition of Key Performance Indicators (KPIs)	2
8.3.2	Frequency	2
8.3.3	Reporting	2
8.3.4	Clinical Data Management System generated, automatic queries	2
8.3.5	Manual queries	2
8.3.6	Follow-up on (persisting) data discrepancies	2

9.	Database closure	2
9.1	Pre-closure data checks	2
9.2	Quality assurance audit and database lock	2
9.3	Database unlock.....	2
10.	Data transfer and exports	2
10.1	Data requests and transfer	2
10.2	Data exports	2
10.3	Export validation	2
10.4	Adverse event data reconciliation	2
11.	Clinical Data Management System archiving and provision of final materials to the sponsor	2
12.	Data preservation	2
13.	FAIR data sharing	2
13.1	Repository	2
13.1	Open Access	2

13.	FAIR data sharing	2
13.1	Repository	2
13.1	Shared artifacts	2
13.2	Data request process	2
13.3	Ethics, legal and security issues	2
13.3.1	Data protection	2
13.3.2	Copyright and intellectual property	2

CTU Bern	Data Management Quality Plan for Study Acronym	Version: xx
	Based on the Template: CS_PM_TEM-10_v01	Page 3 of 28

(Data) Sharing

- Trial protocol
- Patient information and consent form
- Data Management Plan
- Statistical Analysis Plan
- Monitoring reports
- Documentation on protocol deviations
- (Narratives?)
- Data dictionary
- Data
- Statistical code
- ...



Who controls access to data?

Governance and data ownership



Thank you
for your attention!

Sven Trelle, CTU Bern

u^b

b
UNIVERSITÄT
BERN

Workshop: Best Practice with Patient (Person) Data

- What is anonymized data?
- What is coded data?

