

u^b

b

**UNIVERSITÄT
BERN**

u^b

Central Data Monitoring (CDM)

Dik Heg, Danielle Wirz, Mamatha Sauermann

CTU Lecture

22.03.2023

Lecture Overview

1. Background

- What is CDM
- CDM in the context of Risk based monitoring
- Databases

2. Procedure of CDM

- Data cleaning (queries)
- Data monitoring (reports)

u^b

Central data monitoring CDM

Definition: Monitoring NOT done on-site but done centrally or remotely, e.g.

- Phone calls with study sites
 - to follow up on action items, discuss study procedures, etc
- Verify some source documents centrally/electronically
 - site delegation logs, IMP accountability logs, IMP storage temperature logs, etc.
- Review the data collected in the electronic study database (eCRF) and post queries to clarify, confirm or correct (implausible) data and typo errors
- According to **Central Data Monitoring Plan**, which can be very extensive or only minimal.

Recommendations

2013 - Food and Drug Administration (FDA)

*“FDA encourages greater use of **centralized monitoring practices**, where appropriate, than has been the case historically, with correspondingly less emphasis on on-site monitoring”*

Ref: Guidance for Industry – Oversight of Clinical Investigations- A Risk-Based Approach to Monitoring (August 2013)

2016 – ICH E6 (R2) Good Clinical Practice (GCP)

*“The sponsor may choose on-site monitoring, a combination of on-site and **centralized monitoring**, or, where justified, centralized monitoring.”*

Ref: ICH-GCP E6 (R2) section 5.18.3

u^b

Gist of these recommendations

Risk based monitoring approach with a combination and a right balance of on-site and central monitoring methods required to achieve good data quality.

u^b

General aims of CDM

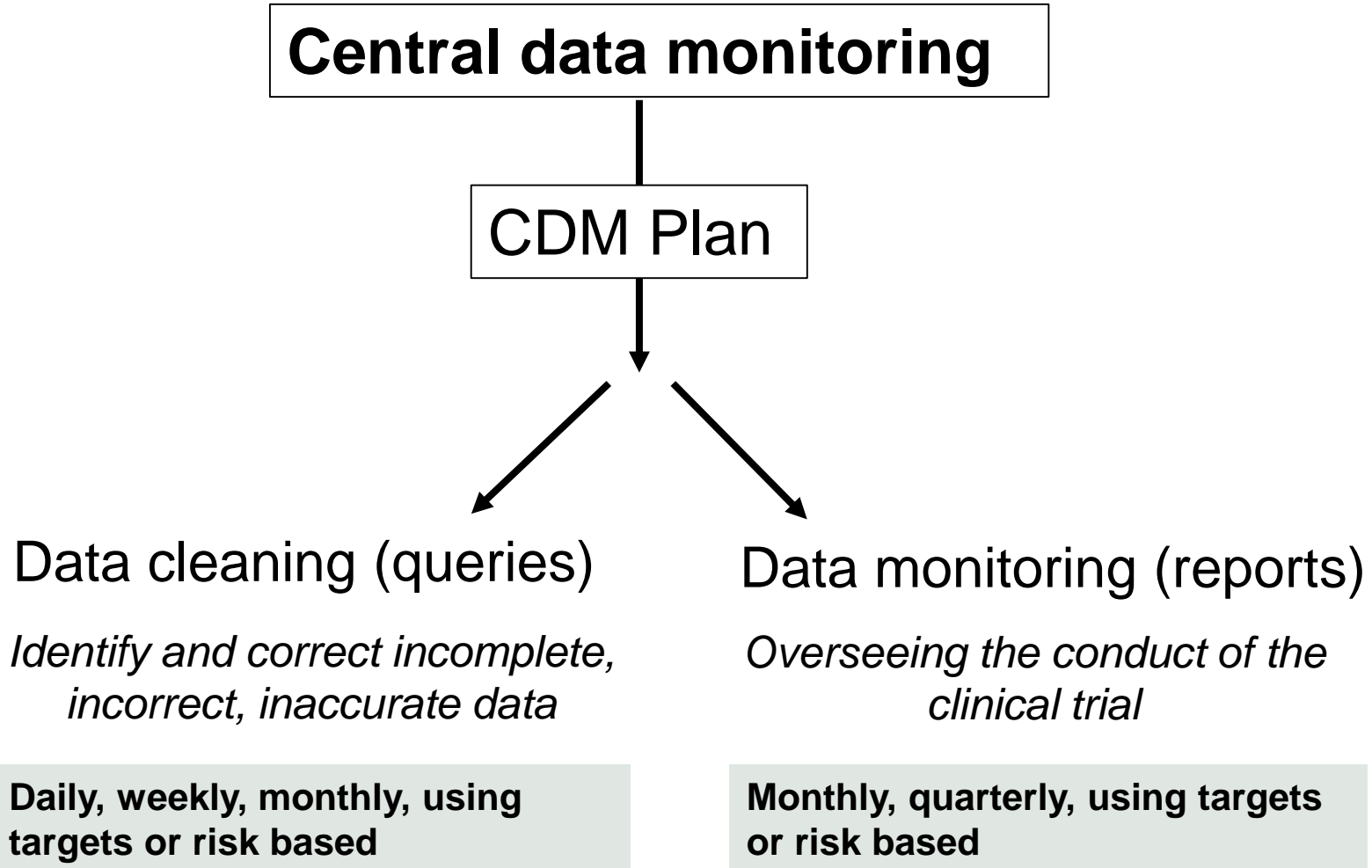
- To enhance the overall **quality** of the trial data – *accurate, reliable, and collected according to the study protocol, GCP rules and other governing regulations.*
- To identify **study related risks and issues** in a timely manner - *ensure the integrity and validity of the trial*
- Extend of CDM depends on **risk assessment of the trial** - *defined inside a Central Data Monitoring Plan*

u^b

Minimal aims of CDM at CTU Bern

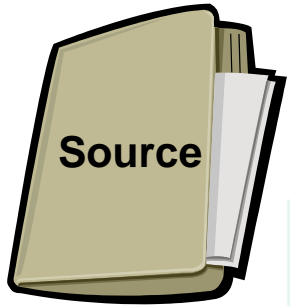
- To enhance the overall **quality** of the primary endpoint and most important secondary endpoints assessments
- To ensure data entry of mandatory data is **complete** (e.g. to describe the population adequately, randomization, adherence, cross-overs, per-protocol population, safety)
- Minimal aims are usually extended according to the risk assessment of the trial and defined inside a **Central Data Monitoring Plan**.

Conduct of CDM



u^b

Data cleaning



Should be between 50 and 250

Height



10

cm

Height



110

cm



Real time checks in the database (**edit checks**)

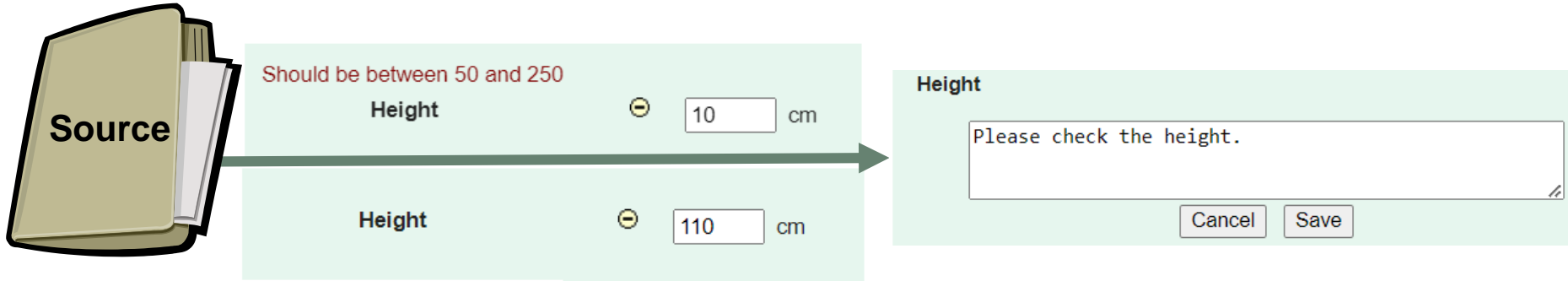
- Outliers (extreme values)
- Missing data
- Transcription errors
- Source check

1st level of data cleaning

Study nurse checks/corrects the data

u^b

Data cleaning



Real time checks in the database (**edit checks**)

- Outliers (extreme values)
- Missing data
- Transcription errors
- Source check

Clarification from study sites (**queries**, email, phone)

- Data completeness
- Plausibility, Typo errors
- Consistency, Accuracy
- Protocol compliance
- Patient safety assessed

1st level of data cleaning

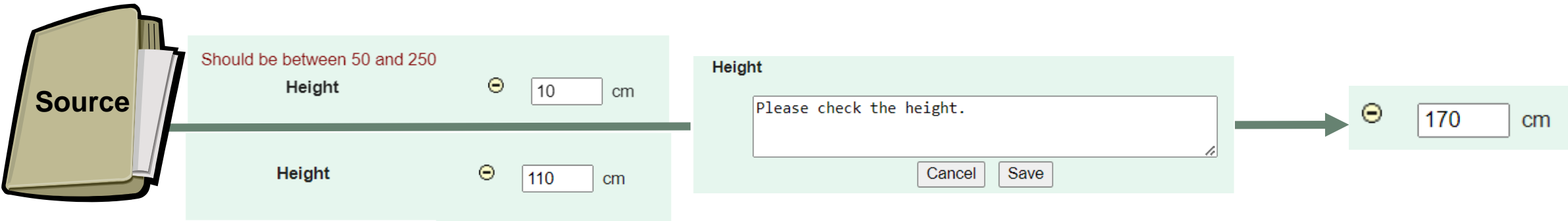
Study nurse checks/corrects the data

2nd level of data cleaning

Central data monitor checks/queries the data

u^b

Data cleaning



Real time checks in the database (**edit checks**)

- Outliers (extreme values)
- Missing data
- Transcription errors
- Source check

Clarification from study sites (**queries**, email, phone)

- Data completeness
- Plausibility, Typo errors
- Consistency, Accuracy
- Protocol compliance
- Patient safety assessed

Statistical analysis trial result/outcome

Final data check (contact CDM, study sites directly)

Quality

1st level of data cleaning

Study nurse checks/corrects the data

2nd level of data cleaning

Central data monitor checks/queries the data

3rd level of data cleaning

Statistician checks

Clinical Data Management Systems (CDMS)

Database, eCRF, queries



ΣSDV	1/2. visit	3. visit (2w)	4. visit/call (3m)	call (6m)	call (9m)	5. visit (12m)	call (15m)	call (18m)	call (21m)	6. visit (24m)	End of Study
🚩	📄	📄	📄	📄	📄	📄	📄	📄	📄	📄	
🚩	📄	📄	📄	📄	📄	📄	📄	📄	📄	📄	
🚩	📄	📄	📄	📄	📄	📄	📄	📄	📄	📄	
🚩	📄	📄	📄	📄	📄	📄	📄	📄	📄	📄	
🚩	📄	📄	📄	📄	📄	📄	📄	📄	📄	📄	
🚩	📄	📄	📄	📄	📄	📄	📄	📄	📄	📄	

- Overview - Patients and Visits
- Overview - Completion status
- Overview - Query status
- Overview - Missings, warnings and errors
- Query details
- Serious Adverse Events (SAEs)
- Random overview
- Protocol Deviation
- CEAC
- CEAC w/procedure date



Record ID	Inclusion	General information	Symptoms	Medication	Diagnosis
1056-1 1	🟢	🟢	🟢	🟢	🟢
1056-2 2	🟢	🟢	🟢	🟢	🟢
1056-3 3	🔴	🔴	🔴	🔴	🔴
1056-4 4	🔴	🔴	🔴	🔴	🔴
1056-5 5	🟢	🟢	🟢	🟢	🟢
1056-6 6	🟢	🔴	🔴	🔴	🔴
1056-7 7	🔴	🔴	🔴	🔴	🔴

REDCap®

🔒 Logged in as dhcg | Log out

📁 My Projects

✉ Contact REDCap administrator

Project Home and Design [-]

🏠 Project Home · 📖 Codebook

📊 Project status: **Production**

Data Collection [-]

📊 Record Status Dashboard

📄 View / Edit Records

Applications [-]

📄 Data Exports, Reports, and Stats

🔍 Data Quality and **Resolve Issues**

u^b

Data cleaning by central data monitor

- **Data completeness** - Are all study visits conducted; are all required assessments and questionings done and is the corresponding data properly recorded in the database?
- **Data consistency** - Are there any inconsistencies or errors in data entry? e.g., if the question “did any new adverse event occur since the last visit?” is answered with “yes”, is there a corresponding event and/or concomitant medication recorded in the database?
- **Data plausibility** - Is the entered data plausible? Pregnant male???



u^b

Data cleaning by central data monitor

- **Protocol and GCP compliance** - Are the required lab values recorded and do the participants fulfill the eligibility criteria; are study-specific visits performed in the correct time window, etc.?
- **Primary and secondary outcomes** - Is data required for the outcomes collected and properly recorded in the database?
- **Participant safety** - Is the informed consent obtained? Are (Serious) Adverse Events recorded and reported in time to the sponsor? Are withdrawal criteria adhered to, etc.? Are safety assessments performed?

u^b

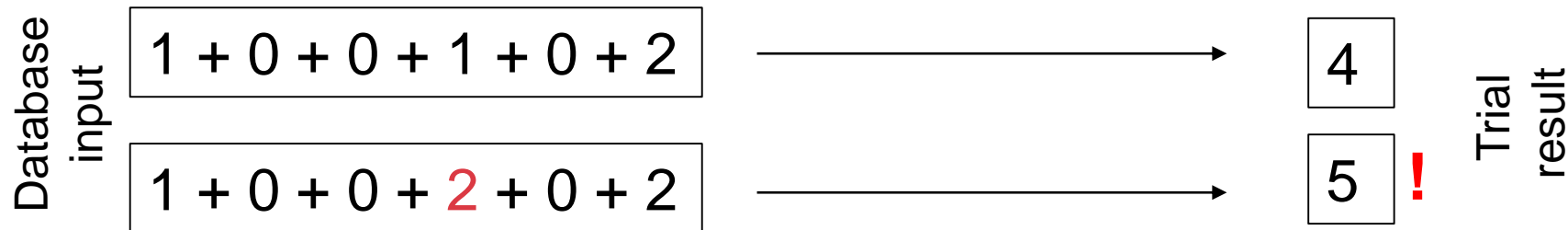
Data cleaning - examples

Incorrect reporting of outcome: e.g. primary outcome (stroke) is reported twice or reporting the same event in 2 different visits

✓ #7428 ? CDMon 11.03.2020 - 16:54 (CET)
The patient was re-admitted to the hospital because of stroke. If you are referring to the same event here again, please change this to 'no'.

! Study nurse 13.03.2020 - 14:38 (CET)
Data updated.

✓ CDMon 31.03.2020 - 15:55 (CEST)
Done.



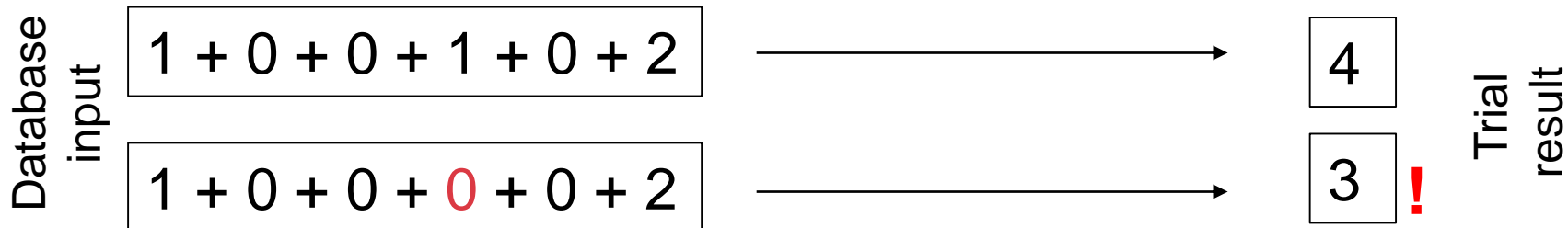
Mistakes can have a direct impact on the trial outcome!

u^b

Data cleaning - examples

Incorrect reporting of outcome: e.g. event does not qualify as a study outcome.

The screenshot shows a data entry form on the left and a query log on the right. The form has several rows with radio buttons for 'yes' and 'no'. The row 'Participant had a clinically relevant non-major bleeding' is highlighted with a black box, and its 'no' radio button is selected. The query log on the right shows a query for 'Non major bleeding' with a result for participant M#18933. The result text states: 'According to Protocol V2.0 the bleeding event only qualifies as a "clinically relevant non-major bleeding" if it fulfills at least one of the following 3 criteria. As this event does not qualify at least one of the criteria, you can change answer to "no".' The log also shows a 'Study nurse' correction on 17.06.2022 and a 'CDMon' completion on 20.06.2022.



Data cleaning - examples

Treatment start/stop: e.g. Treatment stop reported on two different dates

? CDMon	01.04.2020 - 17:23 (CEST)
Thank you. Was the decision to stop the Apixaban already a treatment decision on the 24.12.2019? Because this would then be inconsistent with the information in SAE #3, which states that Apixaban was stopped on the 31.12.2019. I think it cannot be stopped twice, right?	
! Study nurse	30.04.2020 - 15:43 (CEST)
Apixaban stopped on the 31/12/19.	
✓ CDMon	30.04.2020 - 16:19 (CEST)
Done.	

Dates are important! If not checked by EDC system during data entry can have up to 5% typo errors

Data cleaning - examples

Changes in medication: e.g. trial medication stop was incorrectly entered in database even though medication was continued

⊕ Specify why no medication diary is available

▽ ✓ M#17992 ? **CDMon** 07.02.2022 - 08:28:03 (CET)
According to Visit 7 the patient stopped trial medication on 11.11.2021 due to SAE.
Did the patient re-start the trial medication?

! **Study nurse** 04.03.2022 - 01:15:50 (CET)
I had entered it by mistake. Edoxaban was started on 10/11/2021 and has been taken since.

✓ **CDMon** 07.03.2022 - 08:40:18 (CET)
Done.

Changes in Medication must be documented accurately

Data cleaning - examples

Measurement units: incorrect units or no/incorrect conversion

The screenshot shows a laboratory information system (LIS) interface. The main window displays a form for entering laboratory values. The 'Hemoglobin' field is circled in red, showing a value of 137 with units (g/L). A pop-up window titled 'DataCapture - 5.6.3.5 Queries and answers (INSE...)' is overlaid on the right side. The pop-up window shows a list of queries, with the 'Hemoglobin symbol' query selected. The pop-up window contains the following text:

Please mind the conversion (see bottom), hemoglobin should be entered in g/L.
 Please enter your reason for withdrawing the query here:
 [Text input field]
 Withdraw query [v] Cancel Save

Below the main form is a 'CONVERSION TABLE' with the following data:

Analyte	conventional	x Factor	= SI	Example
Creatinine	mg/dL	x 88.4	= μmol/L	1mg/dL is equivalent to 88.4 μmol/L
Glucose	mg/dL	x 0.0555	= mmol/L	70 mg/dL is equivalent to 3.89 mmol/L
Hemoglobin	g/dL	x 10	= g/L	14.5 g/dL is equivalent to 145 g/L
Platelet count	1/μL	x 0.001	= G/L	350'000/μL is equivalent to 350 G/L

If GFR is not automatically calculated, please use the following link:
 * https://www.qxmd.com/calculate/calculator_251/egfr-using-ckd-epi

Hemoglobin 137 g/L = 13.7 g/dL

Double check the correctness of the values that are entered (additional units needed?)

u^b

Data cleaning - examples

Serious adverse event (SAE) reporting:

SAE AND INVESTIGATOR INFORMATION	
Date of initial report	25.07.2021 dd.mm.yyyy
Start date of SAE	15.02.2021 dd.mm.yyyy
Date of awareness of the SAE by the trial site	10.03.2021 dd.mm.yyyy





















SAE not reported to the Sponsor within 24 hours of awareness e.g. in a drug trial

Adhere strictly to SAE reporting timelines!

u^b

Data cleaning - examples

Language must be English – especially for safety reporting

	Specify AE		Need for medical/surgical intervention	
1.	V.a. chronisch-fistelnde Os	 	<input checked="" type="radio"/> yes <input type="radio"/> no	 
2.	hypovolämischer Schock	 	<input checked="" type="radio"/> yes <input type="radio"/> no	 
3.	Hypertension	 	<input checked="" type="radio"/> yes <input type="radio"/> no	 
4.	Plusbilanz Niere	 	<input checked="" type="radio"/> yes <input type="radio"/> no	 
5.		 	<input type="radio"/> yes <input type="radio"/> no	 

u^b

Nr of queries

Trial	Nr of patients	Nr of visits	Nr of queries	Queries/patient
SERVE	100	4	983	9.8
CLEVER ACS	150	2	2048	13.7
PACMAN	300	9	2187	7.3
EVOPACS	308	3	1722	5.6
SCOPE	739	4	10242	13.9
BIOSTEMI	1300	3	2395	1.8
MASTER DAPT	5204*	6	54048	10.4

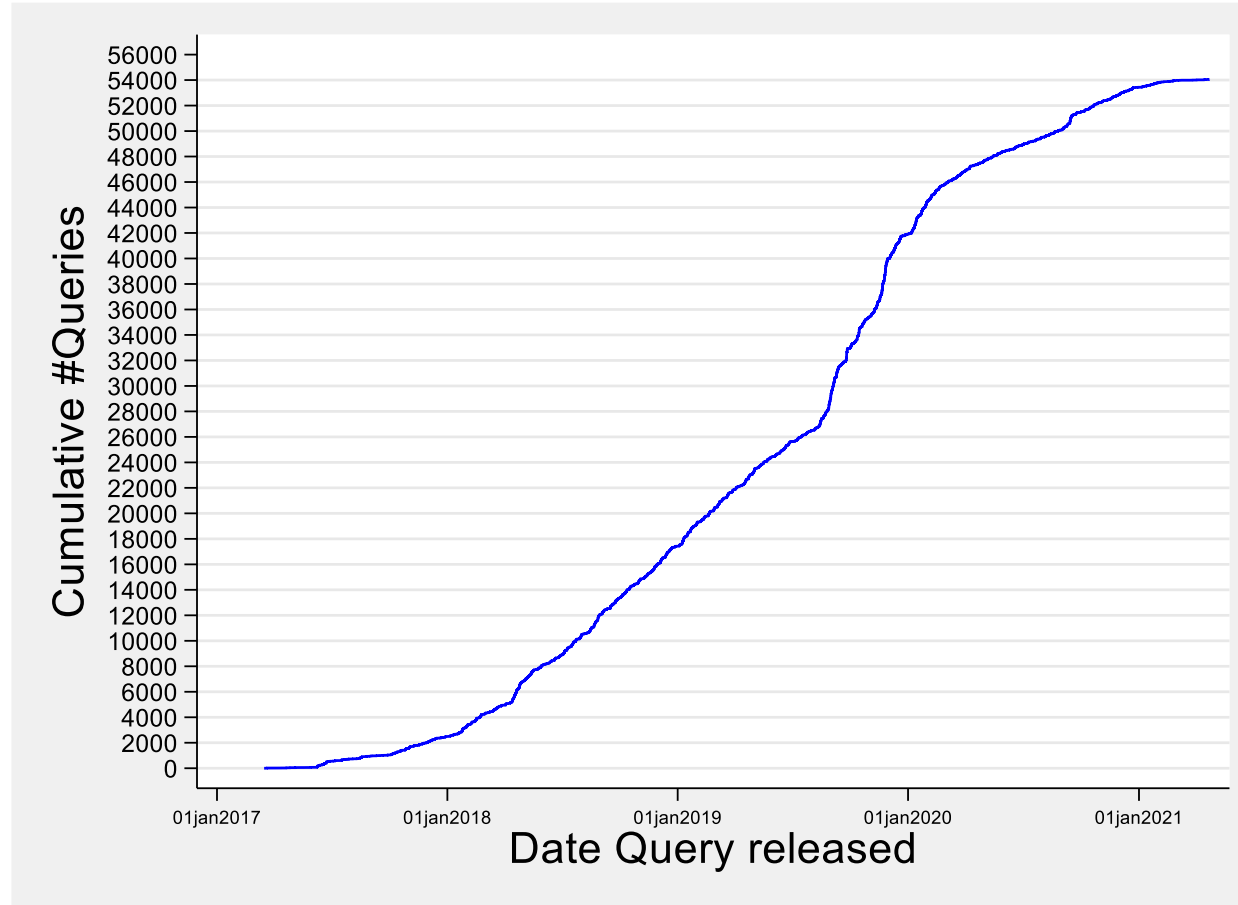
Nr of visits: only counting baseline and all mandatory follow-up visits

*includes n=625 consented but non-randomized patients with only 2 visits

Includes **avoidable queries** due to changes/change of interpretation in the eCRF which need to be queried to complete/amend!

u^b

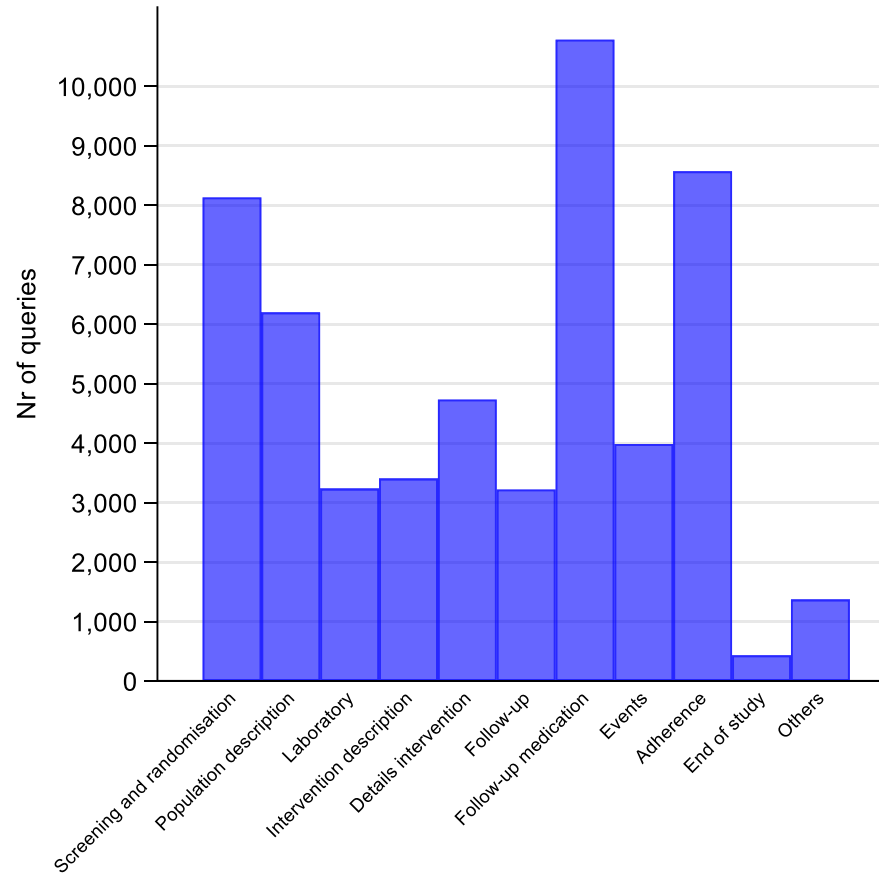
Nr of queries



Unanswered queries and **delayed data entry** are often a problem to close the trial... phone calls, emails...

u^b

Nr of queries



e.g. daily Adherence checked with cross-sectional Follow-up medication = majority of queries

u^b

Data monitoring



Closer view

On-site monitoring

Perspective



Bird's eye view

Central data monitoring

Patterns or trends can be identified only if data is seen as a whole

u^b



VS.



On-site monitoring	Central data monitoring
Visits to the trial site	Data centrally monitored
Source data verification – Review documentation of informed consent, eligibility criteria; ISF/TMF; medical records to assess AEs, SAEs, protocol deviations; Investigational product accountability, storage, etc.	No access to source documents
Personal contact with the site staff – assess site’s familiarity with protocol, involvement of the investigators and PI in the study; verification of proper study conduct , safety of the study participants, etc.	Contact through queries, emails, phone

u^b



VS.



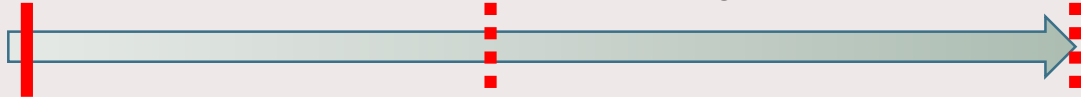
On-site monitoring

Central data monitoring

Study Data

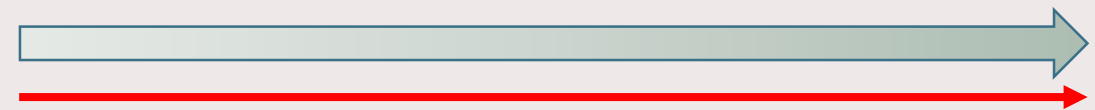
Scattered view of the data

- Data of one site at a time
- Few patients at a time (1-5)
- Data at the time point of monitoring is checked



Overall view of data – data export

- **all trial sites** from the study
- **all patients** included in the study
- **all visits** – FPFV to LPLV



Cross-site trends

Difficult to compare across sites

Identify **cross-site trends**, inconsistencies, and differences between study sites **in real time** –

- Site performance
- Participant safety
- Adherence to ethical principles and participant rights
- Possible data fabrication

u^b

Cross-site trends

Site performance:

Sites with

1. a high number of screening failures, withdrawals, or discontinuations
2. too many or too few protocol deviations
3. a large amount of missing data, missing visits
4. a high number of queries or queries remaining unanswered for a long time, or answers are unclear.
5. Not completing the eCRF in a timely manner
6. Other protocol specific indicators

Patient safety:

1. sites reporting an unusual number of (S)AEs, i.e. considerably more or fewer (S)AEs than the average site
2. many subjects enrolled though all eligibility criteria are not met, or enrolling patients before confirming all eligibility criteria (e.g., patient randomized before checking a blood value that is important for eligibility)

Cross-site trends

Adherence to ethical principles and participant rights:

1. Study-specific examinations or interventions are performed before informed consent was obtained.

Possible data fabrication:

1. Lack of variability or preference of digits in the recorded data across visits of a single subject or across multiple subjects.
 2. Data too perfect
-
- ✓ Export of data from the database into simple tools like Microsoft Excel can be sufficient
 - ✓ Assessment of more complex issues might require statistical methods, e.g. KPI reports (critical performance indicators)

Data monitoring – spot the difference

Study XYZ

Universitätsspital XYZ

Recruitment Report

2.2 Number of samples

Samples from patients with pending definitive consent or which are screening failures or refused to participate are not shown in this table

Compare the occurrence of endpoints across sites:

	Total no. of EVD and LD samples	Total no. of EVD samples	Total no. of LD samples	Total no. of CSF infections*
Site 1	336	293	43	0
Site 2	160	132	28	3
Site 3	27	27	0	1
Site 4	5	5	0	0
Site 5	21	21	0	5
Total	549	478	71	9

% of CSF infections

0.00%

0.02%

0.04%

0.00%

0.20%

*Positive CSF culture without clinical symptoms is not counted as infection in this table.

u^b

Data monitoring – spot the difference

Site 1

	Visits	Protocol deviatio	SAEs
ial			
ial			
ial			
ial			
ial			
ial			
ial			
ial			
ial			
ial			
ial			
ial			

~1% of patients

Site 2

ia			
ia			
ia			
ia			
ia			
ia			
ia			
ia			
ia			
ia			
ia			
ia			

Many SAEs
45.0% of patients

u^b

Data monitoring – spot the difference

	Visits	Informed Consent	End of study	Protocol deviatio
Site 1	001			
	002			
	004			
	003			
	005			
	007			
	008			
	009			
	010			
	012			
	013			
	014			
	016			
	Site 2	01		
02				
03				
04				
05				
06				
07				
08				
10				
11				
12				
13				

Too many protocol deviations

78.5% patients

20.0% patients

Data monitoring report

Central Data Monitoring Progress Report

Study acronym or No.:	
Sponsor:	
Date of report:	
Database:	Choose an item.

Status of central data monitoring (CDM)

No. of subjects planned:	
No. of subjects <enrolled/randomized, recruited etc., based on the study> so far:	
% of planned subjects <enrolled/randomized, recruited etc., based on the study> so far:	%
No. of CRF pages planned to be monitored (total in the study): (Excluding repetitive pages such as <insert repetitive pages i.e., forms that need not be filled in for all participants e.g., (Serious) adverse event, protocol violation forms, etc.>) Note that numbers only refer to CRF pages that need to be monitored according to the monitoring plan.	
Total no. of CRF pages filled in so far (planned and repetitive pages)	
No. of CRF pages with CDM completed (<locked / review A performed / checked>)	
% of CRF pages with CDM completed (<locked / review A performed / checked>)	%
No. of CRF pages with queries (and not <locked / review A performed / checked>)	
% of CRF pages with queries (and not <locked / review A performed / checked>)	%
Note that numbers only refer to CRF pages that need to be monitored according to the monitoring plan.	
Total no. of queries raised so far:	
No. of queries resolved so far:	
No. of queries still open:	
Remarks: <e.g. if any additional forms are checked; if any fields or pages other than mentioned in the monitoring plan were reviewed, etc.>	

Central data monitoring observations

1. General issues
<report here any issues that concern the study as a whole, such as: - Inconsistencies in the study protocol - Any patterns or trends observed in data - Delay in filling in the data in the database - Difficulties in communication with the sites in general, etc.>
2. Database issues
<report here any major issues related to the database i.e., issues with the database itself, and not issues related to the recorded data: E.g. issues that have a major impact on data quality, such as inconsistencies in forms related to end point data, safety reporting forms, implementation of new forms or fields based on CDM findings etc., or any other issues considered important by the CDMonitor.>

u^b

Questions?