# Monitoring site performance in clinical trials

**Dik Heg**

Clinical Trials Unit Bern
Universität Bern
Switzerland

# Risk-based assessment of clinical trials

UNIVERSITÄT BERN

- A large number of clinical trials are aborted prematurely

  e.g. Fogel 2018 Contemporary Clinical Trials Communications

# Risk-based assessment of clinical trials

UNIVERSITÄT BERN

- A large number of clinical trials are aborted prematurely

- Pre-trial assessment of the risk

  RISKLICK (spin-off of CTU Bern) https://www.risklick.ch/

# Risk-based assessment of clinical trials

UNIVERSITÄT BERN

- A large number of clinical trials are aborted prematurely

- Pre-trial assessment of the risk

- Pre-trial survey per site

  Detailed questionaire send to each site

  e.g. expected recruitment, availability of study coordinator, competing trials activity, availability devices

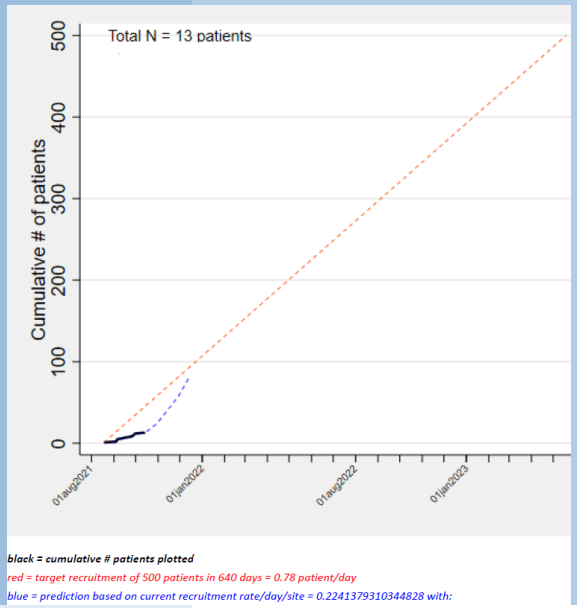
# Risk-based assessment of clinical trials

UNIVERSITÄT BERN

- A large number of clinical trials are aborted prematurely
- Pre-trial assessment of the risk
- Pre-trial survey per site
- **Assessment of the risk during the conduct of the trial**

  e.g.

  Recruitment rate too low (e.g. number of enrolled patients/month)
  Attrition rate too high (e.g. % lost to follow-up or withdrawn consent)
  Primary outcome / Efficacy rate too low (e.g. incidence of non-recurrence)
  Safety outcome rate too high (e.g. incidence of death)
  Protocol deviations too high (e.g. cross-over to non-randomised device)



# What is monitoring site performance?

UNIVERSITÄT BERN

- Sponsor receives regular reports of the quality of the performance of the trial; and separately for each site (using **K**ey **P**erformance **I**ndicators)

- **KPI**s should inform the Sponsor on the progress of the trial and each site; and allow to take **action** to reduce the risks / improve the performance

- Sponsor should always see the *minimum* nr of **KPI**s – which in my opinion are for all trials:

  **Recruitment rate**

  **Attrition rate**

  **Primary outcome / Efficacy rate or assessed**

  **Safety outcome rate or assessed**

  **Major Protocol deviations**

# When to perform more elaborate monitoring of site performance?

UNIVERSITÄT BERN

- Multicentre trials / International trials
  - e.g. differences in standard care
  - e.g. differences in when and how to perform assessments
  - e.g. difficulties in understanding and therefore executing protocol
- Complex trials
  - e.g. complex design (many arms, many dosages, delayed randomisation)
  - e.g. complex pre-screening assessments needed to check for eligibility
  - e.g. cross-over after x days or after x days failure to improve; escalation trials etc.
  - e.g. difficult/new procedure, device or product, etc.
  - e.g. many assessments mandated, many follow-ups mandated, etc.
  - e.g. uncertain or high expected drop-out rates depending on the arm (e.g. high dosage)
  - e.g. training needed on new device
- Complex primary outcome/safety
  - e.g. imaging outcomes, difficult assessment(s), combined (composite) endpoint, calibrated outcomes; many safety assessments mandated

# How to perform this monitoring of site performance?

UNIVERSITÄT BERN

- **Define KPIs**

- Use **Benchmark KPIs** as much as possible usually derived from the Protocol:

| | |
|---|---|
| **Recruitment rate** | e.g. 500 patients/640 days = **0.78 patient / day** |
| **Attrition rate** | e.g. assuming 20% **no MRI** performed at 30 days visit |
| **Primary outcome / Efficacy rate** | e.g. power calculation was performed assuming 10% in placebo group and 20% in verum group = overall **15% efficacy rate**) |
| **Safety outcome rate** | e.g. all patients alive and not withdrawn consent should have ALT measured (**100% in nr of patients measured**)<br>e.g. ALT is expected to slightly increase in all patients on verum, but not above 55 (**% above 60 in nr of patient measured**) |
| **Major protocol violation/deviation rate** | e.g. no cross-overs to other device (**% in patients randomised**) |

# How to perform this monitoring of site performance?

UNIVERSITÄT BERN

Other (flexible) **benchmark KPIs** can be derived using e.g.:

| | |
|---|---|
| **logic/previous trials** | e.g. recruitment **4 patients** / month in large hospitals<br>**1 patient** / month in small hospitals<br>e.g. 60% females and 40% males usually have this condition |
| **benchmark site*** | e.g. 80% MRI performed at Inselspital is the minimum target to achieve<br>(flag sites red if <80%) |
| **average or median*** | e.g. 70% MRI performed overall<br>(flag sites red if <70%) |
| **running average or median*** | e.g. 20% of the patients enrolled in the last two months have high risk score<br>(flag sites red if <20%) |

*disadvantages of these types of KPIs: (1) only be derived if some data of the first enrolled patients are entered;
(2) at least one site is performing well (!); and (3) a single benchmark (e.g. <80%) is often not helpful, so we usually then use e.g.

**red below 25% interquartile**
**orange 25%-50%**
**lightgreen 50%-75%**
**green above 75% interquartile**
etc. etc.

**red = very low**
**orange = too low**
**lightgreen = ok**
**green = very good**

# How to perform this monitoring of site performance?

UNIVERSITÄT BERN

Other **KPIs without a clear benchmark** to cover other risks, e.g.:

**speed to resolve queries***

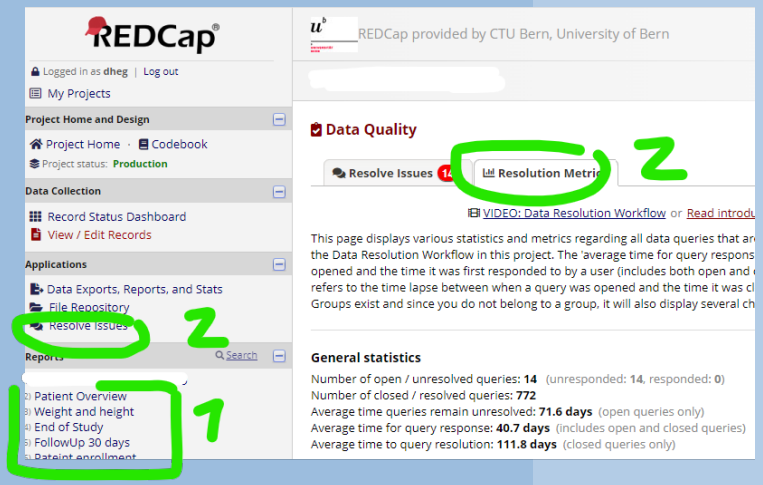
**nr and % of protocol deviations***

**nr and % of cross-overs***

**nr and % of allergic reactions***

**nr and % of surgical / interventional complications***

**etc. etc.**




*disadvantage: these KPIs can only be derived if some data of the first enrolled patients are entered. As there is not a clear benchmark we often provide tables or provide heatmaps:

**red (bad)** to **green (good)**

# Who defines the KPIs?
# Who takes action?

UNIVERSITÄT BERN

Defines the KPIs:

- Sponsor
- Study coordinator
- Project Manager
- Monitor

Statisticians, Central data monitors and Study nurses usually do not have enough knowledge to help here…. CTU Bern Project Managers and Monitors can assist.

Sponsor needs to decide **who** will, *when* and what **action** is taken depending on what KPI(s) are breached:

e.g. **Monitor visit** to site to discuss the *many protocol violations (33% out of six patients)* and **re-train personnel**

It is important to define who, when, what – as these actions need to be documented. It can also be an escalation (e.g. Central Data Monitor calls site, if no improvement after next KPI report, Sponsor calls site). Due to language issues it can be helpful to have the country CRO take action in case of multinational trials.

In conjunction, you will also need to discuss who should receive these KPI reports (Sponsor, Steering Committee, Project Manager, country CRO, Monitors…..) and whether you need to have separate reports per recipient (e.g. country CRO).

# Which KPIs to define?

UNIVERSITÄT BERN

I found it extremely helpful to use PICO :

**PICO**

**P**opulation
experimental **I**ntervention
**C**ontrol intervention
**O**utcome(s) (efficacy and safety)

# Which KPIs to define?

UNIVERSITÄT BERN

## Population

-inclusion/exclusion criteria

(e.g. some sites have many protocol violations; fail to perform/perform too few pre-screening tests)

-risk factors

(e.g. some sites refuse to recruit high risk patients)

-general protocol violations/deviations

(e.g. nr of missed visits; nr lost to follow-up)

KPIs of the **Population** are mostly defined using the complete population (experimental plus Control) – as they are usually based on pre-randomisation/baseline parameters.

# Which KPIs to define?

UNIVERSITÄT BERN

## Device studies

-correct training on the (new experimental) device(s)

(e.g. some sites use operators which are not sufficiently trained on the new experimental device)

-correct use of the (experimental) device(s)

(e.g. some sites have many cross-overs, implant the (experimental) device incorrect, abort too many procedures and surgeries)

-correct use of concomitant procedures

(e.g. some sites prepare procedure not tailored to experimental device, use wrong ancillary devices, do not perform required post-procedure checks)

-correct use of diagnostic requirements

(e.g. some sites refuse/fail to do extra diagnostics needed for the experimental device)

# Which KPIs to define?

UNIVERSITÄT BERN

## Medical product studies IMP

-correct training on the (new experimental) drug(s)

(e.g. some sites do not have all products correctly stored, correct dosages etc.)

-correct use of the drug(s)

(e.g. some sites use wrong dosage, wrong SID/BID, wrong starting point or delayed start, take wrong IMP leading to cross-over; do not perform required first intake checks like immediate allergy)

-correct use of concomitant medications

(e.g. some sites do not up-titrate or down-titrate or do not use the correct concomitant medications; provide concomitant medications which are not allowed in combination with experimental drug)

-correct adherence (per arm if unblinded)

(e.g. some sites have low adherence, too early or too late stops, too many returned pills)

# Which KPIs to define?

UNIVERSITÄT BERN

## Treatment strategies studies

-correct training on the strategy/strategies

(e.g. some sites mix two strategies up; use the incorrect strategy)

-correct execution of the strategy/strategies

(e.g. started/stopped strategies too early or too late; mandated switches not executed; cross-over not allowed and too high)

# Which KPIs to define?

UNIVERSITÄT BERN

## Outcome

-efficacy outcome in overall population (per arm if unblinded)
(e.g. some sites have systematic underreporting, skip too many follow-ups)

-safety outcome in overall population (per arm if unblinded)
(e.g. some sites do not perform post-procedural checks and assessments)

-**Devices**: device deficiency and malfunctions per device
(e.g. some sites have high rates of problems with the experimental device; too low procedural/surgical success)

-**IMP**: (expected/unexpected) side-effects (incl. allergies) in overall population
(e.g. higher than expected rates of side-effects - per arm if unblinded study)

# How to present KPIs?

UNIVERSITÄT BERN

## Tables

- particularly helpful if there are clear **actions** defined!

## Figures/graphs

- particularly helpful if there are no clear benchmarks and therefore there are also no *a priori* actions defined!

# Examples

| Site | Number of patients | First randomization | Last randomization | Patients/month | Follow-up: Performed | Dead | Refused | Excluded | Pending | Imaging performed | Attrition | IMP compliant (≥ X IMPs) | IMP compliance (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 173 | 10.05.201 | 07.10.202 | 4.17 | 129 | 1 | 10 | 1 | 32 | 129 | 6.94 | 127 | 98.45 |
| B | 16 | 25.08.201 | 21.08.202 | 0.44 | 11 | 1 | 3 | 0 | 1 | 11 | 25.00 | 11 | 100.00 |
| C | 15 | 15.08.201 | 03.03.202 | 0.48 | 8 | 0 | 4 | 1 | 2 | 8 | 33.33 | 6 | 75.00 |
| D | 6 | 22.05.201 | 12.12.201 | 0.88 | 5 | 0 | 1 | 0 | 0 | 5 | 16.67 | 4 | 80.00 |
| E | 10 | 04.07.201 | 27.11.201 | 0.59 | 6 | 0 | 2 | 0 | 2 | 5 | 20.00 | 5 | 83.33 |
| F | 34 | 14.03.201 | 22.09.202 | 1.11 | 27 | 0 | 3 | 0 | 4 | 27 | 8.82 | 26 | 96.30 |
| G | 38 | 14.01.201 | 24.08.202 | 1.20 | 30 | 1 | 4 | 0 | 3 | 30 | 13.16 | 28 | 93.33 |
| H | 6 | 14.05.201 | 17.09.202 | 0.37 | 5 | 0 | 0 | 0 | 1 | 5 | 0.00 | 5 | 100.00 |
| I | 2 | 26.02.202 | 26.06.202 | 0.50 | 0 | 0 | 1 | 0 | 1 | 0 | 50.00 | 0 | |
| **Total** | **300** | | | | **221** | **3** | **28** | **2** | **46** | **220** | **11.00** | **212** | **91** |

12.992 Attrition rate (% of patients with performed fup)

## Standard set-up of a table showing each site with e.g.:

**Recruitment rate (Patients/month)**

**Attrition (nr and % of patients with imaging not performed)**

**Adherence (nr and % of patients compliant with IMP)**

**you can add important KPIs to the right of this Table as needed**

# Examples

| | | Risk description | | | | | |
|---|---|---|---|---|---|---|---|
| Domain number | Domain | Definition/explanation | Task frequency | Task complexity | Risk Consequence | Risk level | Original data source |
| **1** | Randomized population | Patients **randomized** fulfill at least one of the inclusion criteria and do not fullfill any of the exclusion criteria at **Randomization** | 5 | 0.4 | 3 | **6** | Medical record |
| **2** | Non-randomized population | Patients **not randomized** fulfill at least one of the inclusion criteria at **Screening**, but do not fulfill any of the inclusion criteria at randomization and/or fullfill at least one exclusion criteria at **Randomization** | 5 | 0.4 | 3 | **6** | Medical record |
| **3** | Population description | The **population characteristics** can be adequately described, and are according to the expectations or are according to the target population, as described in the protocol | 2 | 0.4 | 1 | **0.8** | Medical record |
| **4** | Randomized treatment | Patients executed the **randomized regimen** | 4 | 1 | 6 | **24** | Medical record |
| **5** | Follow-up | Patients assessed up to the **15 months follow-up alive**, or at **death** before 15 months | 4 | 0.4 | 3 | **4.8** | Medical record |
| **6** | Primary endpoints | **Primary endpoints** adequately assessed up to 15 months follow-up (or up to death if within 15 months) | 4 | 1 | 6 | **24** | Medical record |
| **7** | Secondary endpoints | **Secondary endpoints** adequately assessed up to 15 months follow-up (or up to death if within 15 months) | 4 | 1 | 3 | **12** | Medical record |
| **8** | Potential SAEs related to randomized regimen duration | **SAEs** for which it cannot be excluded that they are related to the randomized regimen, are adequately assessed up to 15 months follow-up (or up to death if within 15 months) | 4 | 0.4 | 3 | **4.8** | Medical record |
| **9** | General protocol compliance | Patients executed the protocol according to the **mandatory items**, and **deviations from recommendations are rare** and explanable from medical and electronic records | 4 | 1 | 6 | **24** | Medical record |
| **10** | Deadlines and funding | The trial is completed within the **deadlines** | 4 | 0.4 | 6 | **9.6** | eCRF |

# Examples

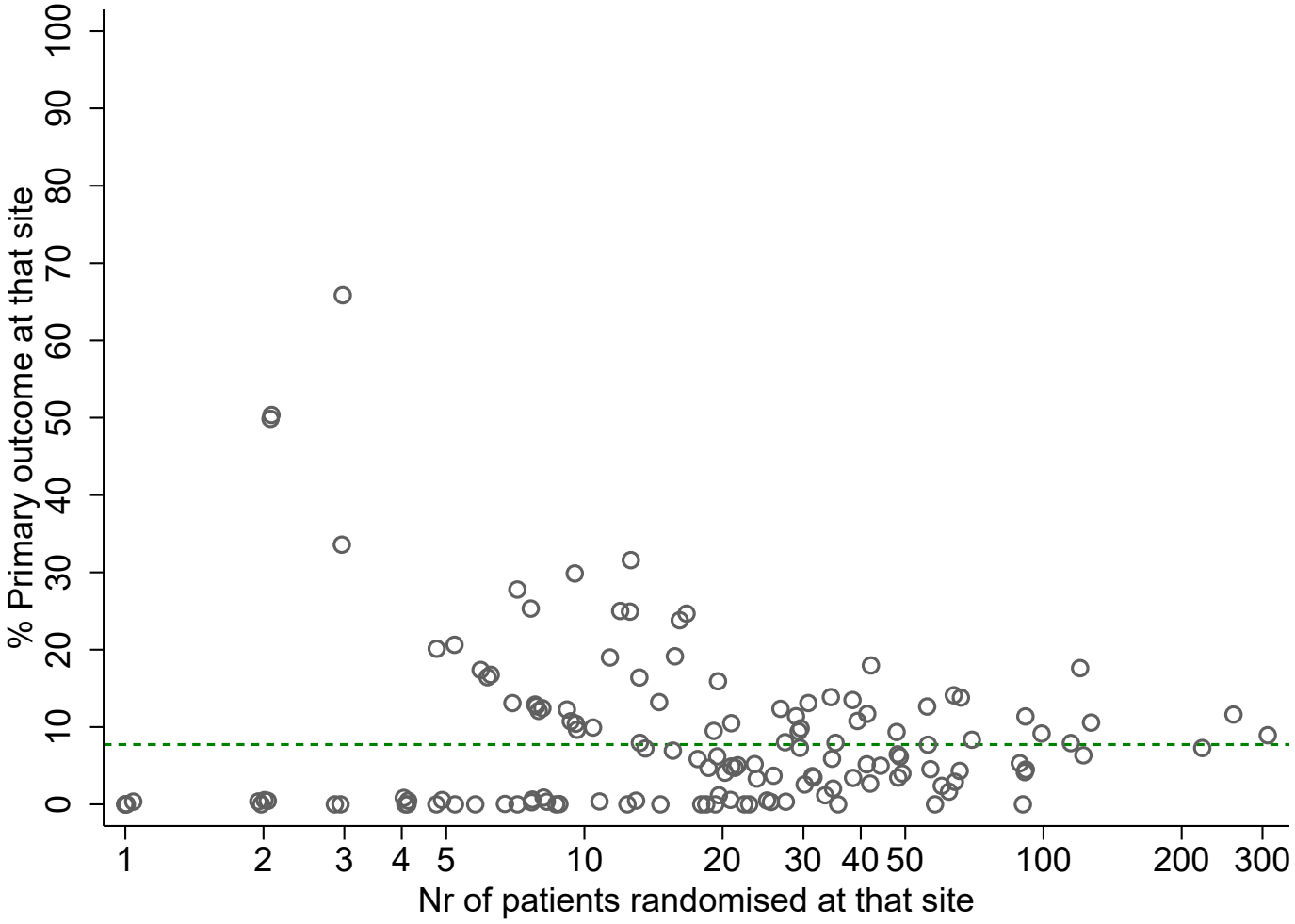
**All sites need to be reminded to use correct follow-up windows (Newsletter)**

**Site J needs retraining in inclusion/exclusion**

**Sites D, I, J, L and M need retraining in using correct randomisation window**

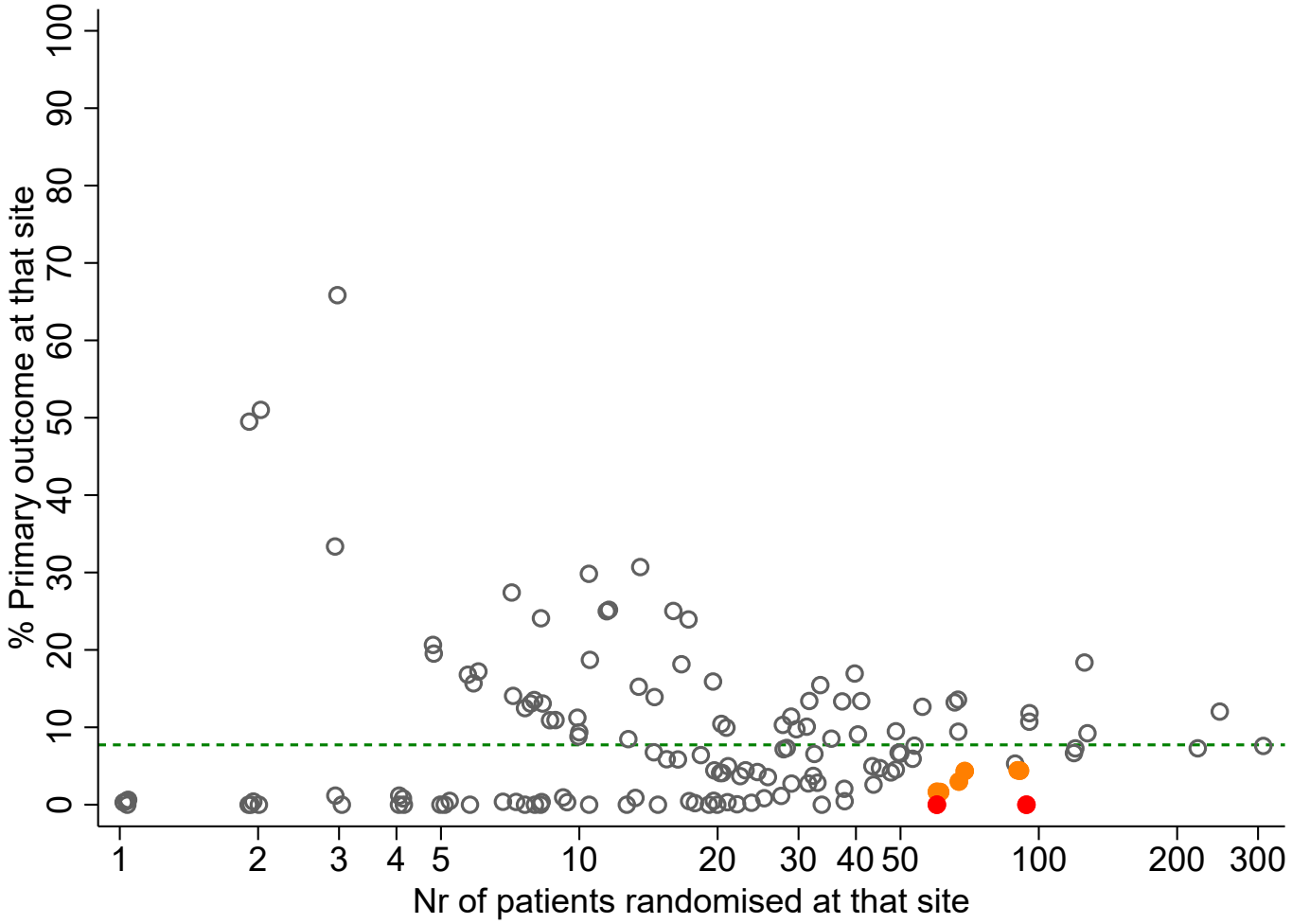
| | | Sample size | | 1 Randomized population | | 3 Population description | | | withdrawal | | | lost to follow-up | | 5 Follow-up timeline | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Randomization exclusion criteria violated | Randomization within correct window violated | Randomization too late | Randomization when not on concomitant medication | Clinical characteristics insufficiently completed | no Intervention | Incorrect device | Patient withdrawal up to M3 | Patient withdrawal up to M6 | Patient withdrawal up to M12 | Patient withdrawal up to M15 | Patient lost to follow-up up to M3 | Patient lost to follow-up up to M6 | Patient lost to follow-up up to M12 | Patient lost to follow-up up to M15 | Patient outside of timeline at M3 | Patient outside of timeline at M6 | Patient outside of timeline at M12 | Patient outside of timeline at M15 |
| | | | Randomized population | <4% | <5% | <5% | <3% | <3% | <1% | <2% | <2% | <3% | <5% | <6% | <2% | <2% | <2% | <2% | <2% | <5% | <2.5% | <10% |
| Country | Site | Total | Rand | 1.02 | 1.03 | 1.04 | 1.05 | 3.01 | 3.02 | 3.03 | 5.01 | 5.02 | 5.03 | 5.04 | 5.05 | 5.06 | 5.07 | 5.08 | 5.09 | 5.10 | 5.11 | 5.12 |
| | All sites | 5204 | 4587 | 1.2 | 5.5 | 0.0 | 0.6 | 0.8 | 0.0 | 0.4 | 0.1 | 0.0 | 0.0 | 0.1 | 0.1 | 0.2 | 0.2 | 0.2 | 7.4 | 8.7 | 10.0 | 9.8 |
| A | A | | | 0.9 | 1.7 | 0.0 | 1.3 | 0.0 | 0.0 | 0.9 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 7.0 | 8.7 | 4.0 | 4.8 |
| A | B | | | 1.9 | 3.7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.9 | 3.8 | 3.9 | 4.0 |
| A | C | | | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 10.5 | 5.6 | 0.0 |
| B | D | | | 0.0 | 33.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 33.3 |
| B | E | | | 0.0 | 4.8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 9.5 | 0.0 | 4.8 |
| C | F | | | 1.1 | 3.4 | 0.0 | 4.6 | 3.4 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.2 | 24.7 | 17.9 |
| C | G | | | 0.0 | 0.0 | 0.0 | 0.0 | 6.9 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 17.2 | 17.9 | 7.1 | 21.4 |
| C | H | | | 1.5 | 4.5 | 0.0 | 3.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 3.0 | 3.0 | 4.5 | 6.1 |
| C | I | | | 0.0 | 15.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 20.0 | 25.0 | 40.0 | 15.0 |
| D | J | | | 40.0 | 20.0 | 0.0 | 0.0 | 20.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 100.0 | 0.0 | 0.0 | 0.0 |
| D | K | | | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 100.0 | 100.0 | 100.0 |
| D | L | | | 0.0 | 6.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 2.1 | 2.2 | 0.0 | 2.3 |
| D | M | | | 0.0 | 12.8 | 0.0 | 0.0 | 2.1 | 0.0 | 2.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 6.4 | 4.3 | 18.2 | 6.8 |

# Examples



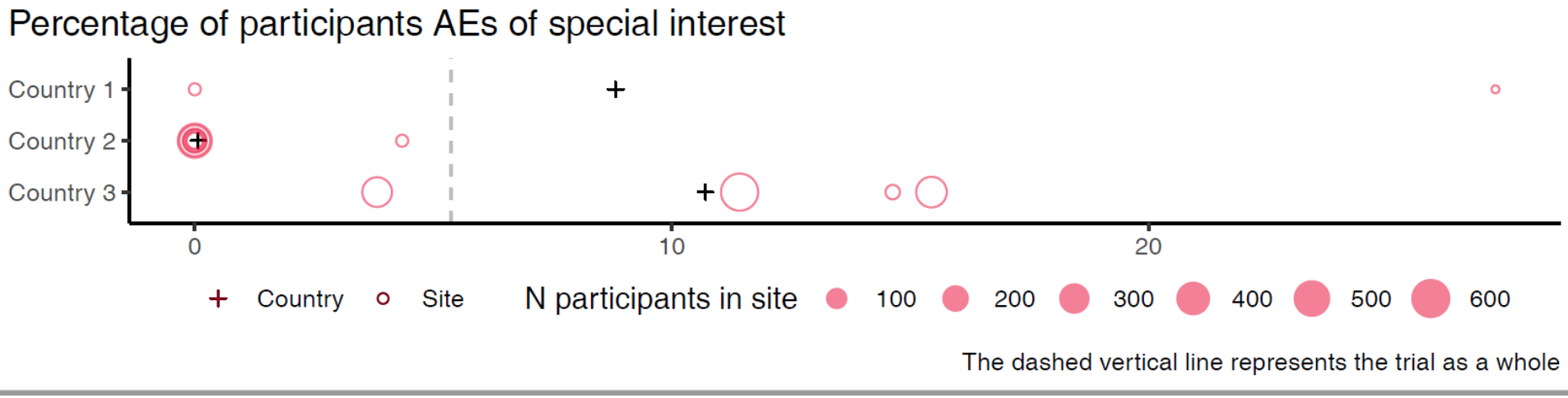

**Which sites would you consider suspicious, i.e. unexpected low rate of the primary endpoint?**

# Examples



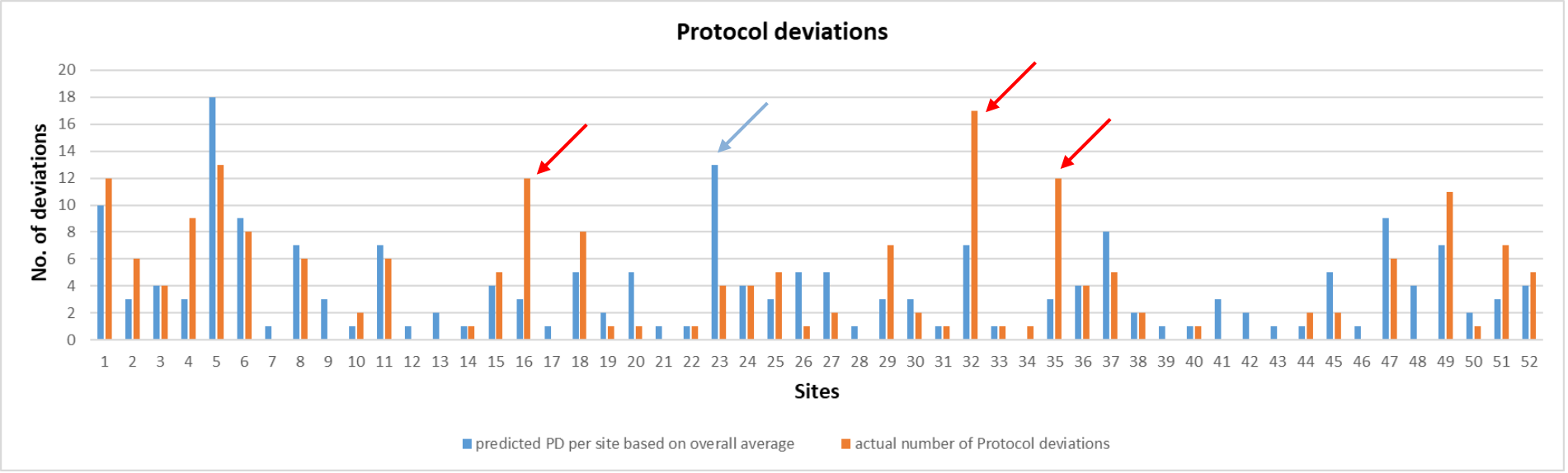

**Half-funnel shape expected – orange/red sites need retraining on how to assess more completely the primary outcome**

# Examples




**Countries 1 and 2 have too low rates of Adverse Events of special interest (retraining in event capture and data entry) – note: only Country 2 has a large nr of patients enrolled at the moment**

# Examples



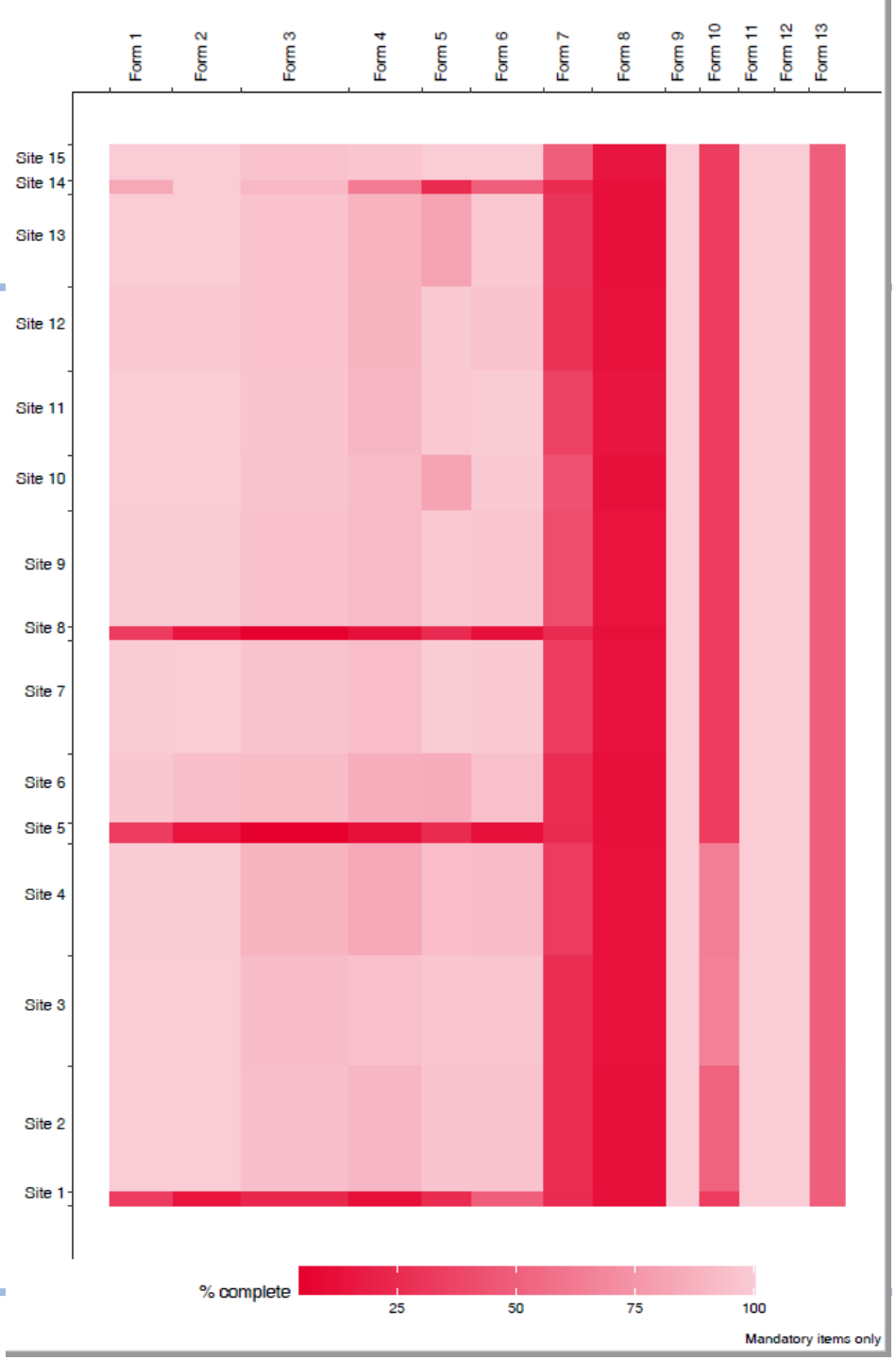

**Many sites have too few Protocol deviations (PD) – in particular site 23 = probable sign of underreporting (add queries for derived PDs using statistical code)**

**Some sites have too many Protocol deviations (PD) – in particular sites 16, 32, 35 = need retraining on the Protocol**

# Examples - heatmap

Forms **1** to **13**
(each mandatory item one column)

Sites **1** to **15**
(each patient one row)



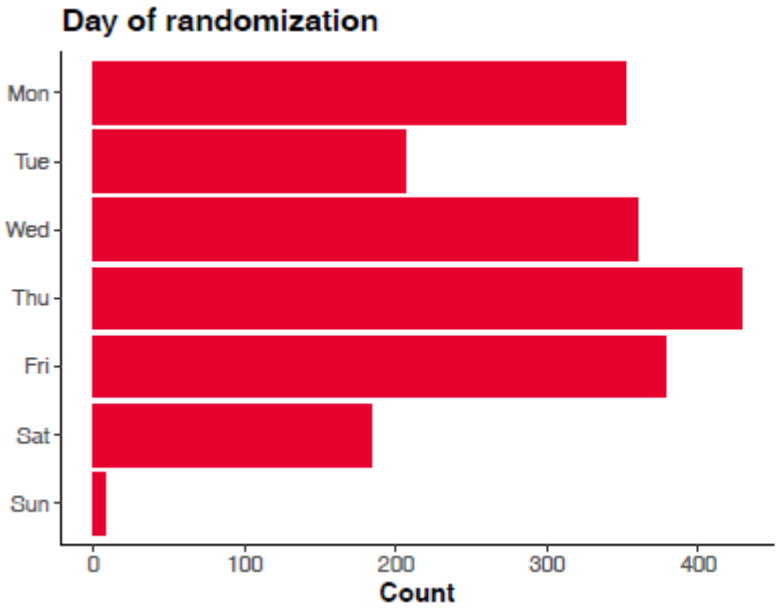

**Sites 1, 5 and 8 need to be contacted by phone or email with reading confirmation (too much delay in data entry)**

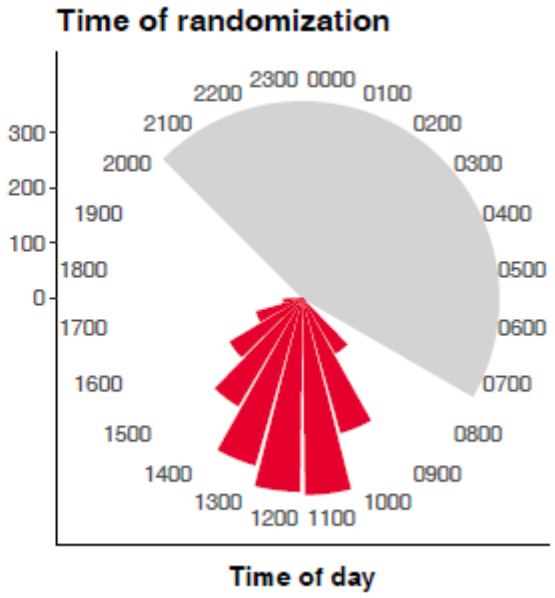
**Forms 7, 8, 10, 13 need a deadline for data entry completeness (Newsletter)**
**(currently missing because some of these forms are from visits planned in the future)**

# Examples

**Elective** procedures during the day only – rarely/never in weekends
**Emergency** procedures also during the weekends and night

This example: randomisations are expected to occur during the day, and can also occur during the weekends, so this example is plausible.



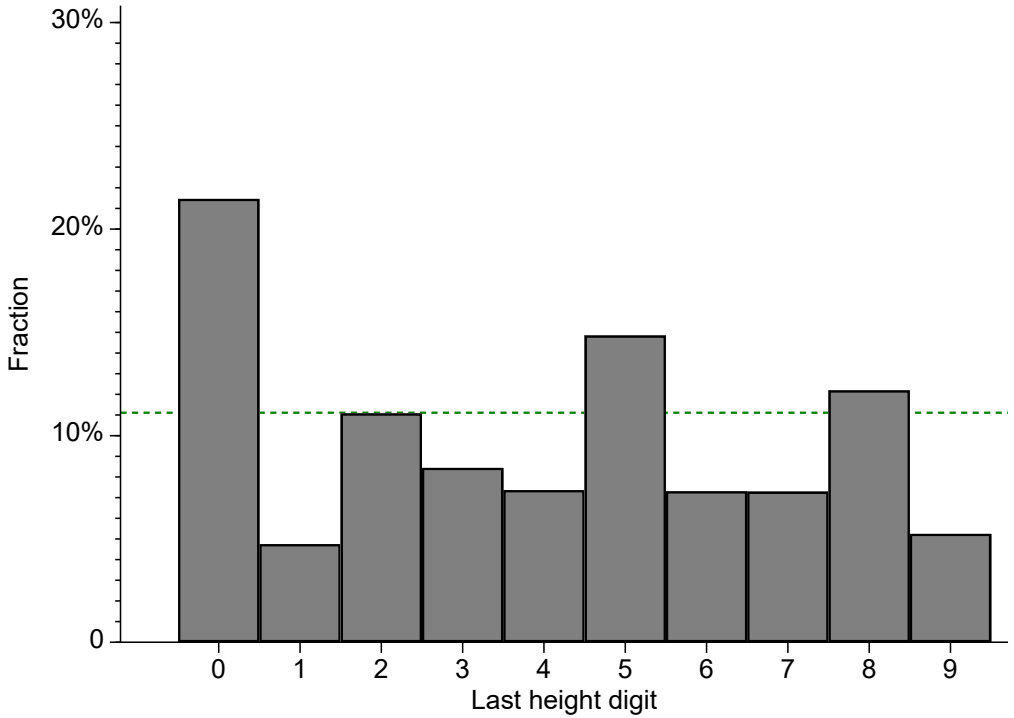
# Examples

UNIVERSITÄT BERN

## **Digit preference** (0 to 9 should occur equally often)

Height as reported by patient:

17**0** cm -> shown is the last digit **0** in the graph:



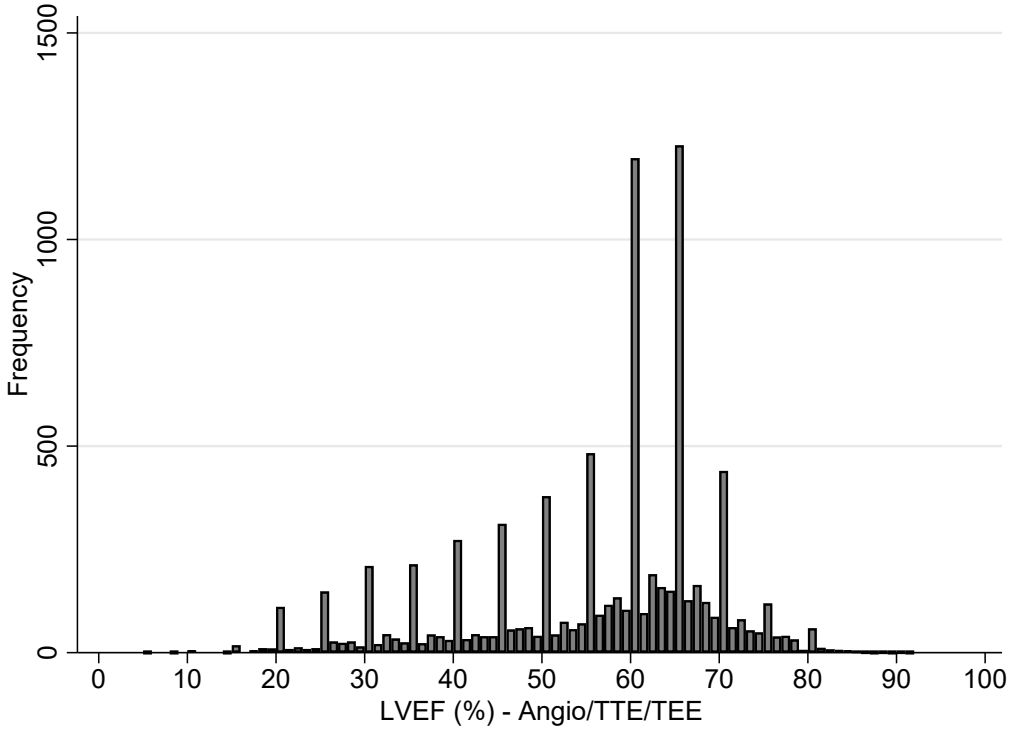

e.g. 17**0** cm as reported by the patient, showing a digit preference to „round“ height 169 and 171 to 170 (and probably also „round“ 3, 4, 6 and 7 into 5).

Conclusion: **re-measure height** during the clinical visit if you need e.g. a precise body surface area measurement (BSA)

# Examples

UNIVERSITÄT BERN

## **Digit preference** (0 to 9 should occur equally often)




e.g. Left ventricular ejection fraction LVEF is usually recorded as 10, 20, 30, 40, 50, 60, 70, 80 etc.; but sometimes with higher precision depending on the methodology, so this example of a digit preference for trailing zeros is plausible.

# Related activities

UNIVERSITÄT BERN

**Data Safety Monitoring Board DSMB**

– receives unblinded reports and advises the Sponsor on e.g. the safety of the patients in the trial.

**Sponsor visit**

– in case of severe breaches of several KPIs the Sponsor can delegate a person to visit the site (**sponsor visit**)

# Related activities by CTU Bern

UNIVERSITÄT BERN

**Data Management**

– make certain all relevant parameters are captured inside the Electronic Data Capture System (REDCap, secuTrial, …)

– separate Protocol deviations eCRF

**Statistics**

– KPI programming and reporting

– Assists Central Data Monitor

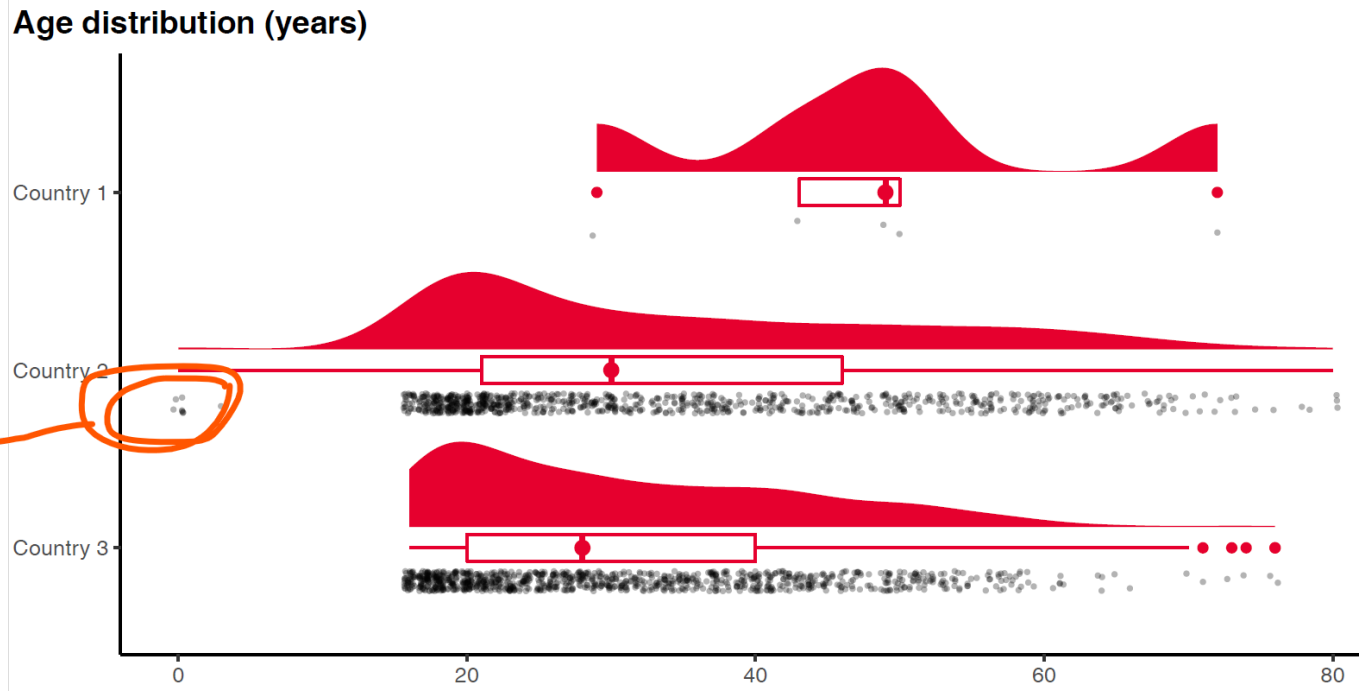
# Related activities by CTU Bern

UNIVERSITÄT BERN

## **On-site** and **Central Data Monitoring**

– reports

- queries




Can also be delegated to take action (extra **on-site visit** or **retraining**)