DATA ANALYSIS AND REPORTING

Met statistiek meer kans

Corine Baljé-Volkers

11 oktober 2019

Getting acquainted

- Corine Baljé-Volkers, head of B&P
- Author! : 1998 MW, alter also STATS.
- All kinds of medical indications,
  multiple large and small clients.
- 16 employees, of which 4 located
  in the north of the Netherlands.

Content

- Use of MDR 217/745, ISO 14155 (GCP), ISO 13485
- Examples

Disclaimer:

- Limited to where ***Statistics*** (think broad!) is involved.
- Main focus on EU; FDA similar in many ways.
- I am assuming some basic statistical knowledge… ?

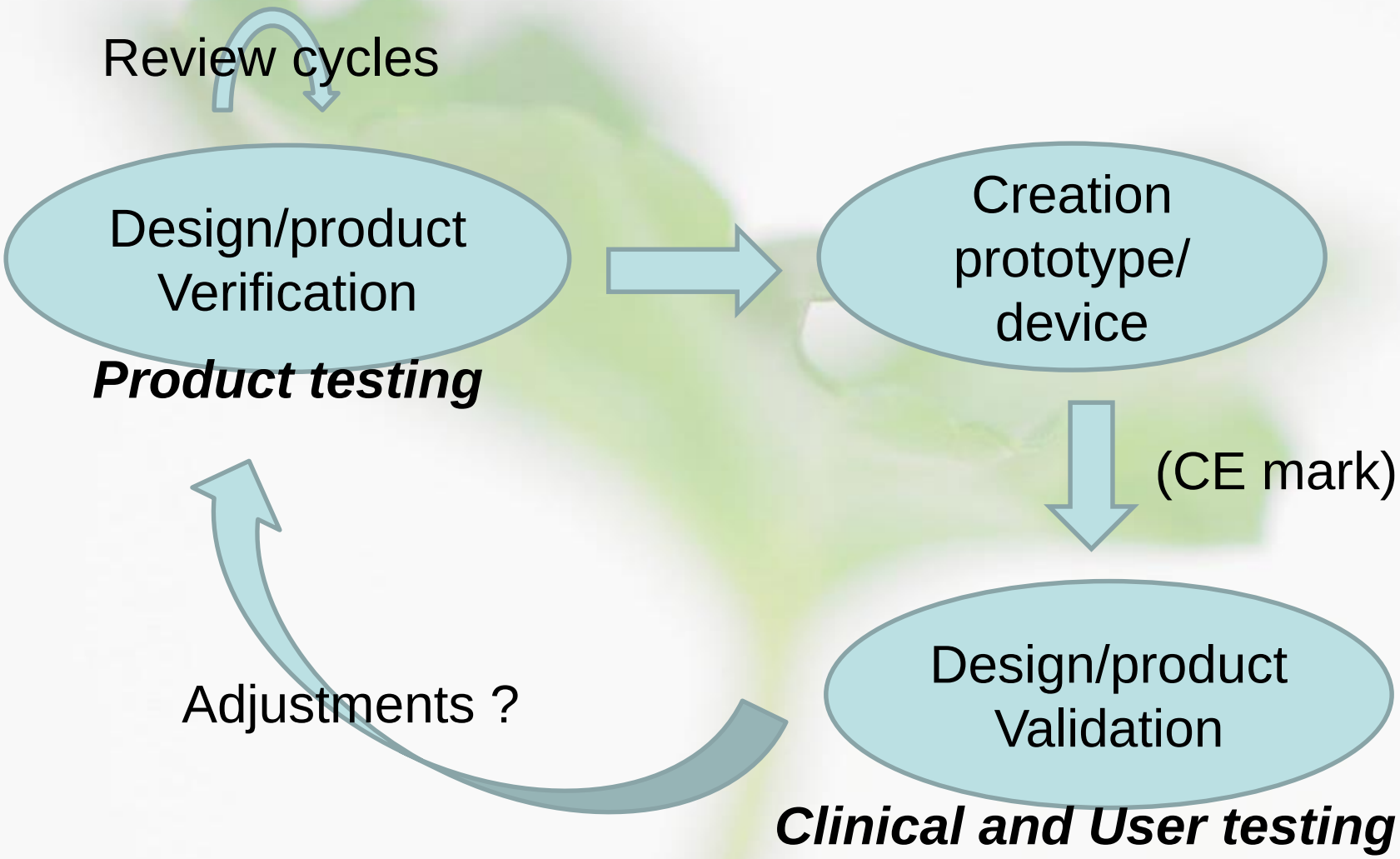
Content

- Today's goal:
  - Change the mindset
  - Learn from each other through discussion(s)

MDR 217/745

- Published in May 2017, effective of May 2020 for MD (2022 for IVD)
- Adhering to the guidances (MDD, AIMDD) = 95% ready for regulation.
- ***Stronger demands on data collection and reporting of clinical data*** → DM, STATS and MW more important
- ***Performance → Effectiveness:*** there is a greater emphasis on comparative device evaluation.
- Refer to ISO 14155, effective Dec 2019 (?)

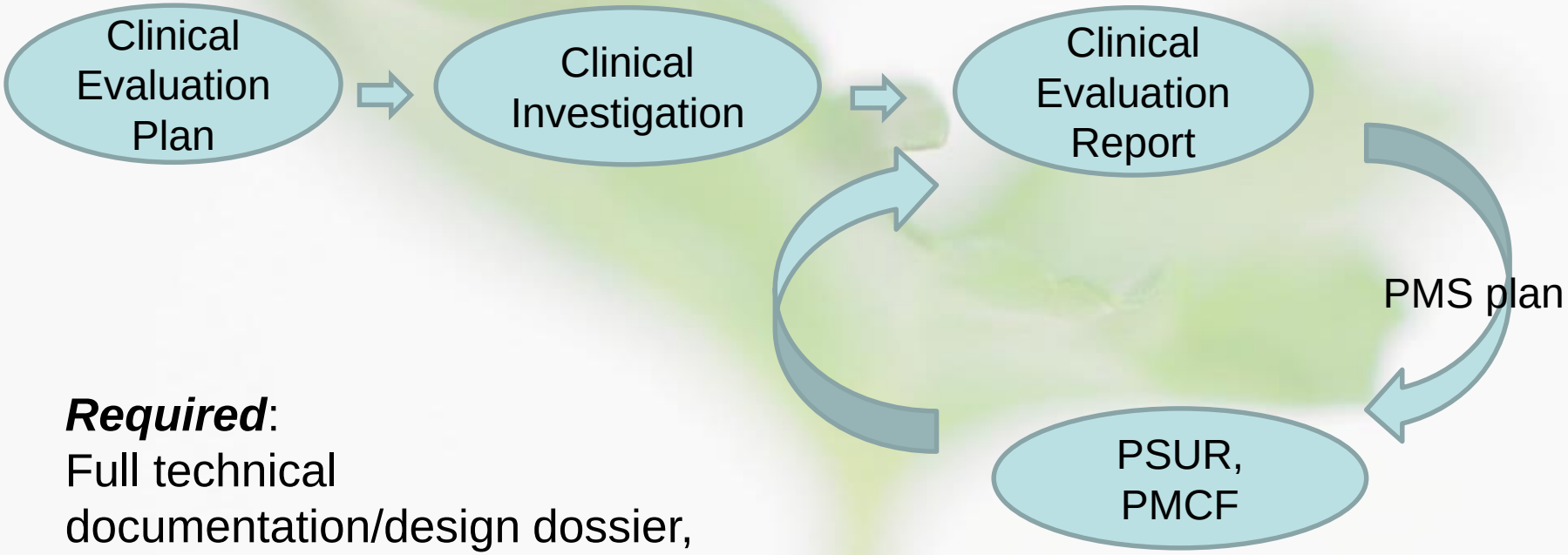
Design/product V&V



V&V

- Design Verification **Plan** is important ➔ define Design Inputs, and how to verify.
- Four main verification methods are inspection, demonstration, testing, and analysis.
- Design Validation **Plan** may involve clinical evaluation (actual or simulated)
- Both require sample size justification (sometimes N=1 is enough).
- Statistical methods applied: process control charts, pass/fail criteria & analysis, specification limits → key term: **variation**.

Life cycle MD



Required:
Full technical documentation/design dossier, incl CER.

Summary in EUDAMED

(Bi-)yearly updates

Required Content of Technical Documentation (per MDR)

Technical Documentation:

1. Device description and specification, including variants and accessories
    - 1.1 Device description and specification
    - 1.2 Reference to previous and similar generations of the device
2. Information to be supplied by the manufacturer
3. Design and manufacturing information
4. General safety and performance requirements
5. Benefit–risk analysis and risk management
6. Product verification and validation
    - 6.1 Pre-clinical and clinical data
    - 6.2 Additional information required in specific cases

Technical Documentation on Post Market Surveillance:

PMS plan (post-market surveillance plan) and report

PSUR (Periodic Safety Update Report)

ISO 14155

DRAFT INTERNATIONAL STANDARD
ISO/DIS 14155

ISO/TC **194**	Secretariat: **DIN**
Voting begins on: **2018-06-19**	Voting terminates on: **2018-09-11**

Clinical investigation of medical devices for human subjects — Good clinical practice

Investigation clinique des dispositifs médicaux pour sujets humains — Bonnes pratiques cliniques

ICS: 11.100.20

➔ *Addresses GCP for design, conduct, recording and reporting of clinical investigations carried out in human subjects to assess safety and performance of medical devices for regulatory purposes.*

ISO 14155

- Most important changes compared to 2011 version:

1) Referring to GCP principles
2) Registration of clinical investigation in publicly accessible database (EUDRAMED)
3) Clinical Quality Management
4) Risk-based monitoring
5) Statistical considerations → Annex A
6) Ethical committees
7) Enforcing risk management during clinical investigation
8) <…>

ISO 14155/GCP

- Investigations shall be scientifically sound and described in a clearly detailed CIP.
- Each individual involved in designing, conducting, recording and reporting a clinical investigation shall be qualified by education, training and experience to perform his or her respective task(s).
- All clinical investigation related information shall be recorded, handled, and stored in a way that allows its accurate reporting, interpretation, monitoring, auditing and verification.

Annex I ISO 14155

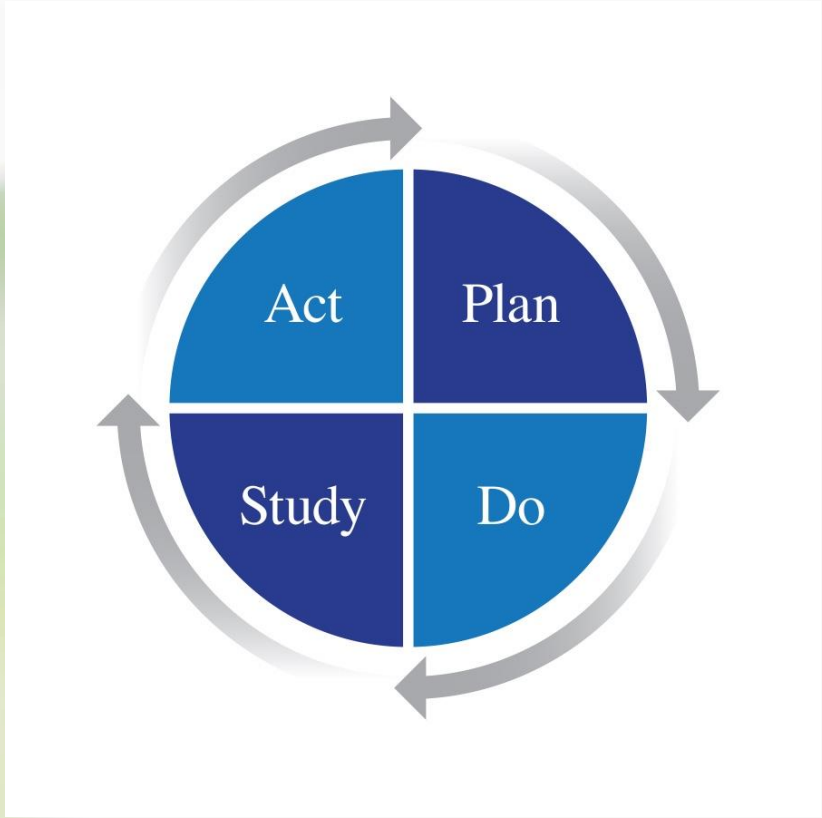
Regulatory status	**PRE-MARKET**		**POST-MARKET**	
Clinical Development stage	Pilot	Pivotal	Post- market	
Type of design	Exploratory/ confirmatory	Confirmatory		Observational
Description clinical investigations	First in human clinical Investigation / Early or traditional feasibility clinical investigation	Pivotal clinical investigation	Post market clinical investigation	Registry / Post market clinical investigation [(a)]
Burden to subject	Interventional			Non-interventional

[(a)] Registry data may be used for pre-market regulatory purposes, this may also apply to the post market clinical investigation data.

General process

Deming Circle !

- **Plan**
- **Do**: Collect information (data/documents/..)
- **Study**: Evaluate information (statistics ?)
- **Act**: Report (or repeat)



Plan

- Clear and detailed document that describes in advance on how you will go about answering the question you want to answer.

- Think about: (main) objective, data/information to collect, test samples, analyses, report. Key items also include: risk management/analysis, sample size, who is your population, etc..

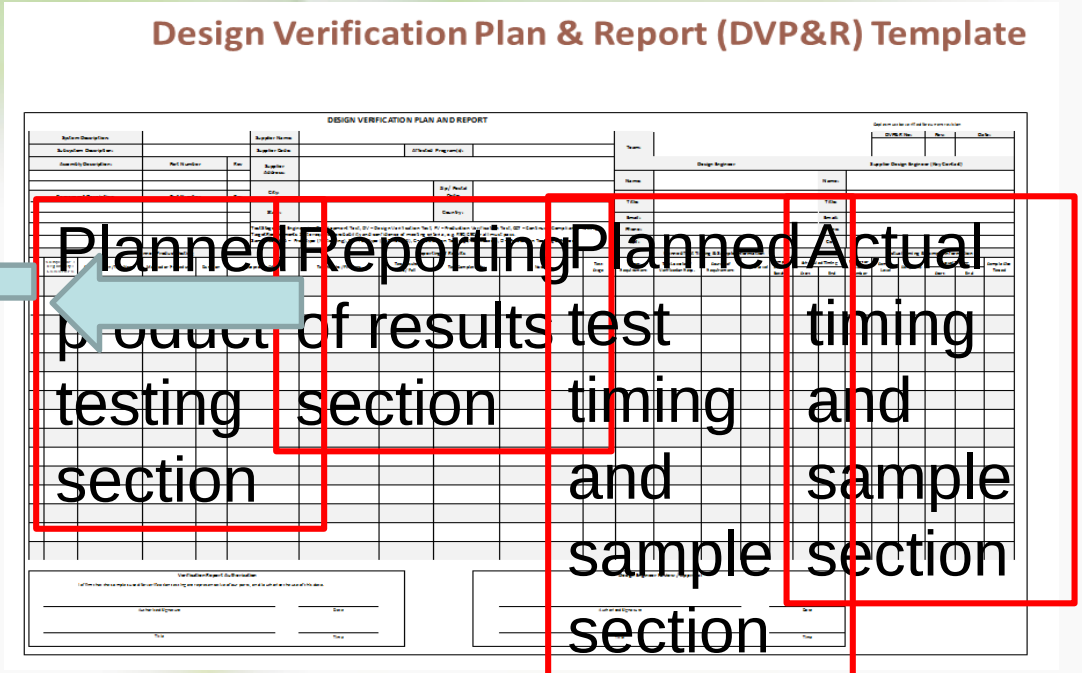
➔ **Make sure you are well prepared for the ride**

Plan

- Plan can be as simple as a table format displayed over 2 pages, or a much more extensive document for e.g. a clinical evaluation in humans.
- Can even be a combination document with reporting included!

Test number, Test stage, compliance, statistical level of test name, confidence method (how) meeting the design duration, criteria, test acceptance location, sample criteria (to pass level, sample the test). For size/batch number, each functional scheduling time component or start/end of test. ingredient.

***Advice**: Make your own standard forms/ documents !*



Plan; ISO 14155 - Annex A: CIP

- Hypotheses/objectives
- Design clinical investigation
- Subjects
- Statistical substantiating
- Data management
- Adverse events

➔ Input from a statistician is necessary.

Plan - Design: Endpoints/hypothesis

Defining efficacy and safety endpoints that are *scientifically measurable* and *objective*.

Hypothesis set:

H_0: null-hypothesis, “situation you’re in”

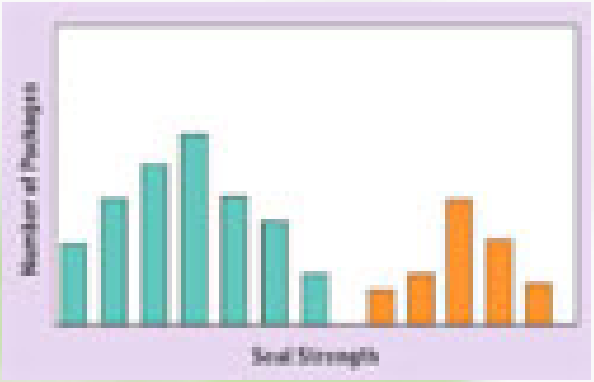
H_A: what you want to prove

Plan - Design: Endpoints/hypothesis

- *Example 1*

Does temperature have an effect on seal strength ?

Packaging sealed at 2 temperatures



➔ Hypothesis set:

H_0: temperature does not have an effect (2 test groups equal)

H_A: temperature does have an effect (2 test groups different)

Plan - Design: Endpoints/hypothesis

- *Example 2*

Evaluating samples to see if a certain level of acceptance has been met. E.g. a packaging defect rate, spec. states < 1%.

➔ Hypothesis set:

H_0: defect rate ≥ 1%

H_A: defect rate < 1%

Plan

- **Take home message**:
  - Create a comprehensive protocol with rationales, design, test method description, data analysis (against acceptence criteria), sample size justification, etc..
  - Make sure, especially for clinical studies, to minimize the number of missing data.

Statistical substantiating: sample size

- Sample size is depending on a lot of factors that need to be pre-specified and substantiated in the plan.
  - Information available from engineering studies/non-clinical performance studies or maybe literature, etc.. → use as basis.
  - Medical device specifics (e.g. type of data, variability in output, risk, ...)
  - Study design, hypothesis, primary objective & analysis, interim, ..
  - What are your units (subjects, runs, samples, ...)
  - What confidence level is specified/desired
  - Maximum allowable error (bias, dispersion)
  - Likelihood of occurrence of a specific event

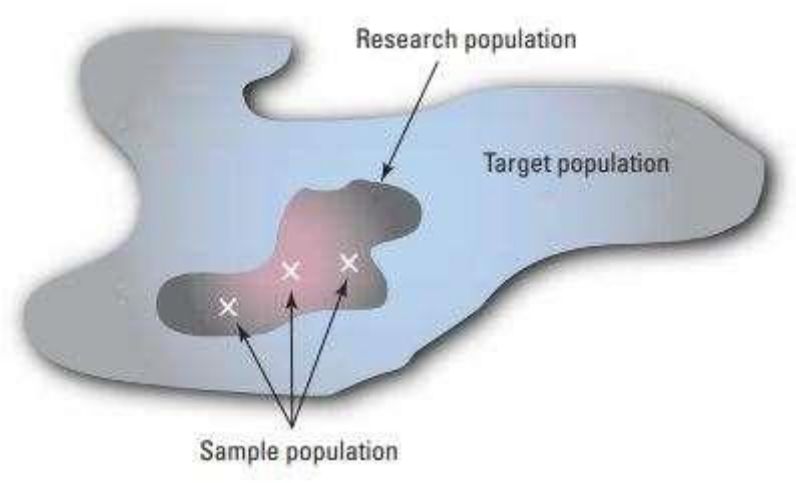
Statistical substantiating: sample size

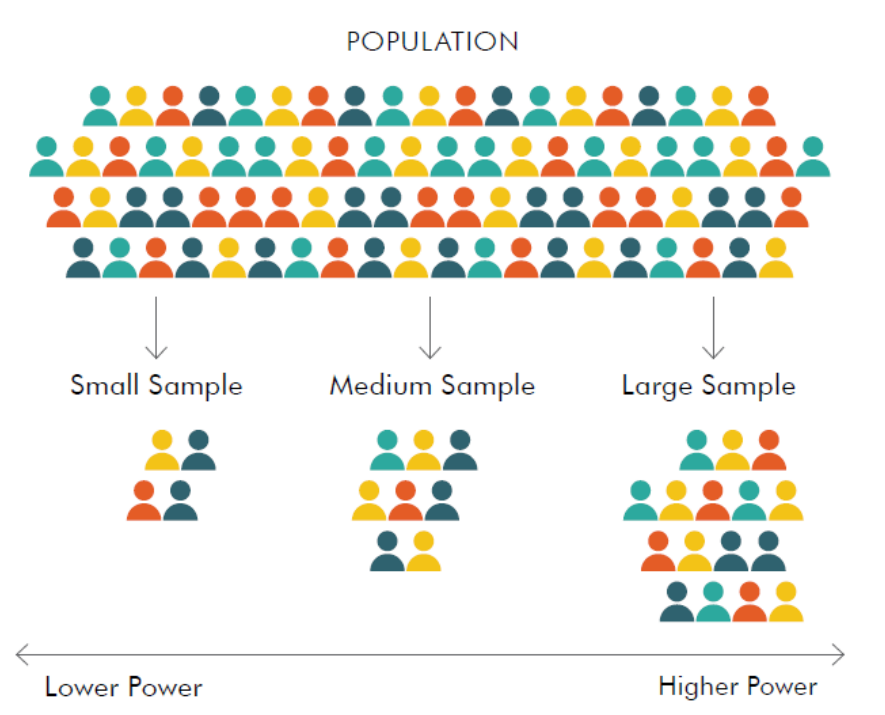
➔ *There is no single sample size method that works for every situation.*

- Even non-statistical possible, as long as you have a **clear rationale**.

- Sample size is impacted by:
  - Variation (more variation = higher sample size)
  - α and power (higher power = higher sample size)
  - Effect size (small effect size = higher sample size)

Statistical substantiating: sample size





Sampling lets you draw conclusions or make inferences about the population or product lot from which the sample is drawn.

Statistical substantiating: sample size

- **Rule of thumbs**. Should **never** be used as rationale, but only for mindset.
  - For every additional factor in your statistical model you roughly need 10-20 units extra in the sample size
  - Statistical analysis can only be applied with a minimum of 20 units
  - For every statistical test, you have a 5% chance to find a significant result by chance, therefore after 20 tests you have at least 1 which is purely chance.
  - Range can be “guestimated” from standard deviation, by multiplying with 4.
  - A CV >= 1 indicates a relatively high variation, while a CV < 1 can be considered low.

Statistical substantiating: sample size

- **True facts**:
  - Categorization of data requires a higher sample size.
  - Some NB prefer to have a sample size of approx 40, even non-statistical.
  - If you are using multiple statistical testing, you should correct for this (α). It will lead to a higher sample size.
  - Preferably use the same statistical model for the primary research question and the sample size, or be more conservative.
  - The alternative hypothesis is what you are trying to prove, not the other way around.
  - Sample size determination is not a push on the button, but needs a lot of thinking: ***The use of mathematical formulas is no substitute for thinking...***

Statistical substantiating: sample size

H_0: defect rate ≥ 1%

H_A: defect rate < 1%

Take:

alpha=5%, power=90%

P_0=0,01

➔N=473 to show a difference between defect rate of 0,1% and 1%

P1	N
0,001	473
0,002	773
0,003	1182
0,004	1693
0,005	2546
0,006	4293
0,007	8094
0,008	19189
0,009	80774

Statistical substantiating: sample size

Proportion non-conforming. Sampling plans are selected based on the confidence statement that can be made if they pass. E.g.: 95% confidence that more than 99% of the units meet the requirement (denoted as 95%/99%), linked to risk. ➔ less then 1% non-conforming

Deciding on the Acceptance Quality Limit will define the sample size (based on risk assessment)

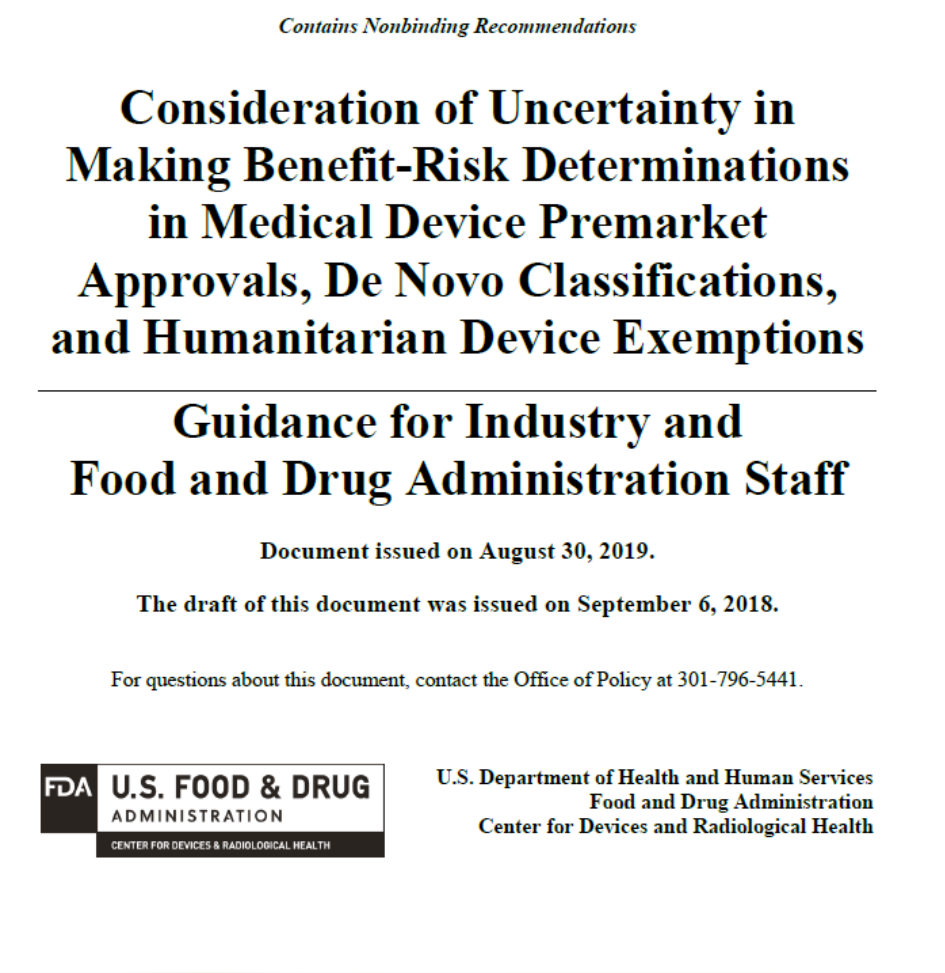
299 samples, 95% confident that the true defect rate is 0-1%.

Risk	Confidence	Reliability	Defect Free Samples Required
High	95%	99%	299
Medium	95%	95%	59
Low	95%	90%	29

Statistical substantiating: sample size

FDA guidance –

30 August 2019



Contains Nonbinding Recommendations

Consideration of Uncertainty in Making Benefit-Risk Determinations in Medical Device Premarket Approvals, De Novo Classifications, and Humanitarian Device Exemptions

Guidance for Industry and Food and Drug Administration Staff

Document issued on August 30, 2019.

The draft of this document was issued on September 6, 2018.

For questions about this document, contact the Office of Policy at 301-796-5441.

FDA U.S. FOOD & DRUG ADMINISTRATION
CENTER FOR DEVICES & RADIOLOGICAL HEALTH

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health

Statistical substantiating: sample size

Breakthrough Device, benefit-risk factors and the relevant non-clinical and/or clinical information ➔ performance goal of 70% of the treated patients experiencing treatment success is acceptable, in a single arm study.

Hypothesis set:

H_0: LCL of success rate < 70%, H_A: LCL of success rate ≥ 70%

Sample sizes dependent upon extent of “uncertainty”.

Statistical substantiating: sample size

Summary: One-sided significance levels and differences in sample size of premarket study

Scenario	One-sided significance level	Sample size†	Postmarket data collection and other measures in light of the greater uncertainty
Case 1: Conventional Uncertainty	2.5%	535	Not applicable
Case 2: Modest Uncertainty, Modest Postmarket Data Collection	5%	385	Modest postmarket data collection as a condition of approval Flag postmarket data collection on FDA's website
Case 3: High Uncertainty, Substantial Postmarket Data Collection	20%	125	Robust postmarket data collection using a registry (or other appropriate mechanism) as a condition of approval If appropriate, inclusion of information about the postmarket data collection and its purpose in labeling as a condition of approval and in the SSED Flag postmarket data collection on FDA's website

† Based on Clopper-Pearson binomial confidence interval. For illustration only, these calculations do not account for statistical power.

Statistical substantiating: sample size

- **Take home message**:
  - Sample size justification is very much “specifics-dependent”
  - Not only controlled by number of units, but also e.g. by number of uses.
  - Need a set that represents the variability in the “entire population”
  - Needs to be related ro risk (low risk, lower sample size)
  - Need to have enough knowledge on factors having an impact.

Data collection

Data collection

Per ISO 14155, clinical data is needed to demonstrate that a device is safe, performs as intended, and has an acceptable risk/benefit ratio.

Generally, in the EU, even Class I devices require clinical evidence demonstrating that the level of device effectiveness consistently and accurately meets requirements for the labeled application. This ***clinical data may include bench testing, a compilation of relevant scientific literature, and/or a clinical trial in human subjects.***

FDA: 21 CFR part 11 ; electronic records and electronic signatures

Data collection

- Currently: mostly use of spreadsheets (excel) for data collection ? And printed document storage ? Not mentioned in plan ?
- Possible issues with data ownership, privacy issues, data or document control.
- Requirement: reliability, integrity, control, traceability ➔ ***traceable and credible***.
- What do you need ?
  - Database system
  - Document management system
  - Procedures

Data collection

- A good system is:
  - Validated
  - All initial data and changes saved (automated audit&edit trail)
  - Prevents unauthorized access
  - Authorization list
  - Can be locked for changes
  - Backup in place
  - Possible randomization/blinding is maintained
  - Identification code for subjects
  - Long-term data storage

Data collection

- Types of data/document collection, e.g.:

Literature study:

- Complete publication references, selection criteria, weight of publications
- Database with all raw “data”
- Literature search report

PMS report: Incidents reports, registries, user feedback (e.g. issues with instructions for use) etc..

Trending data: AEs, complaints, recalls, defect rates/product failures, …

Self-generated clinical data for assessing performance

Data collection

- Will provide you with:
  - Effective acquisition of the data/documents (+ sharing)
  - Traceability
  - Privacy/confidentiality maintained
  - Easy and controlled monitoring/cleaning of data
  - Possibility for data integration (structured !)
  - Easier data presentations, statistical analyses

Data collection

Specifically for **FDA**: supplying data in CDISC standardized format.

- SDTMIG for Medical Devices v1.1, Release Date: 1 Feb 2019

https://www.cdisc.org/standards/foundational/sdtm-ig-md

Defines data standards for medical device-related data in clinical research. SDTMIG-MD v1.1 covers devices under investigation as well as ancillary devices in non-device trials and is appropriate for regulatory submission and non-submission studies.

Data collection

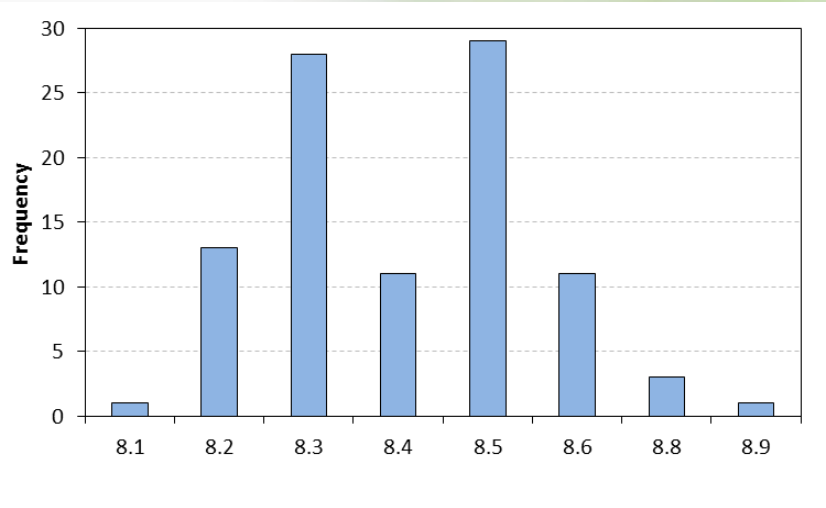
- **Take home message**:
  - Always make sure you collect the right data to answer your question (plan !)
  - Quality: are the data sufficiently valid to support a conclusion?

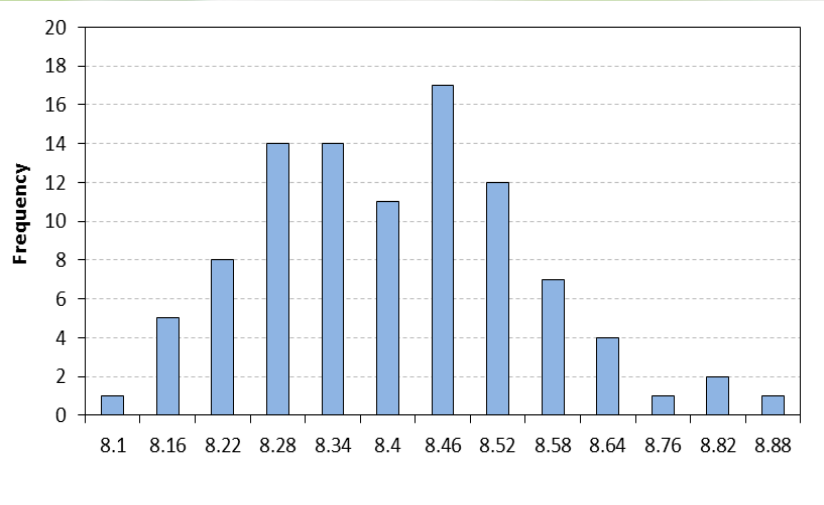
Data presentations/results/report

- Statistical evaluations have become more important.
- Can be limited to descriptive or including statistical modelling.
- Primary analysis should be in line with sample size estimation method.
- Cleary plan the statistical analyses *before* having access to the data (can be in plan, or separate statistical analysis document).
- Preferably have the statistical analysis performed by someone independent to avoid bias.
- Get to know your data before doing any analysis.

Data presentations/results/report

- Example (of not knowing your data).





➔ Means similar (8,4), large difference in variation (0,20 versus 0,15).

➔ It was therefore recommended to report and use the results with 2 decimal places for both higher accuracy and precision,

Data presentations/results/report

Several Fitbits versus counting actual steps (‘golden standard’).

Several circumstances.

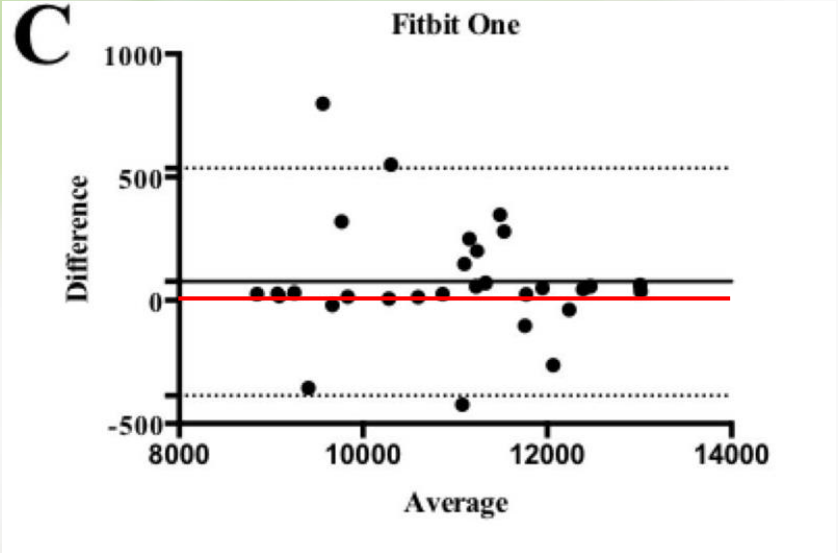
The level of agreement between 2 devices was examined using the Bland-Altman plot.

Total step counts:

Actual: 10950 ± 1209

Fitbit: 10875 ± 1236

Underestimating, although not statistically significant.



Data presentations/results/report

- Reporting should be complete and fully transparent: **what** did you do, **why**&**when** did you do it, did you **deviate** from plan (and if yes, why), etc..
- First describe results without conclusion.
- Add conclusion/interpretation/discussion regarding the results.

Your conclusion/discussion should:

- Answer the research question(s) and restate the major findings.
- Clearly state the “flaws” that may impact the results.
- Provide reasonable argumentation (if possible, combined with literature research on the specific subject)

Data presentations/results/report

- **Take home message**:
  - Know *all* of your data.
  - Be critical, fair, transparent !

Final sheet

<u>Main message</u>:

Ensure scientific conduct of the investigation and credibility of the investigation results in *every aspect of the MD development*

- THANK YOU -