Zur Rolle der Terminologie bei Interimanalysen

Lektionen aus der COVID-19-Pandemie

Benjamin Hofner
Head of Data Science & Methods

Bundesinstitut für Impfstoffe und biomedizinische Arzneimittel
Federal Institute for Vaccines and Biomedicines

Das Paul-Ehrlich-Institut ist ein Bundesinstitut im Geschäftsbereich des Bundesministeriums für Gesundheit.

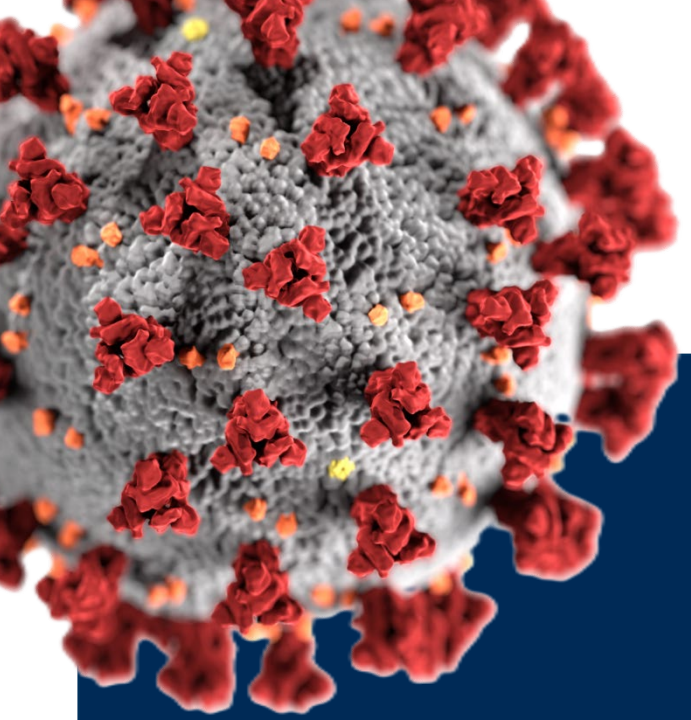
The Paul-Ehrlich-Institut is an Agency of the German Federal Ministry of Health.

Paul-Ehrlich-Institut

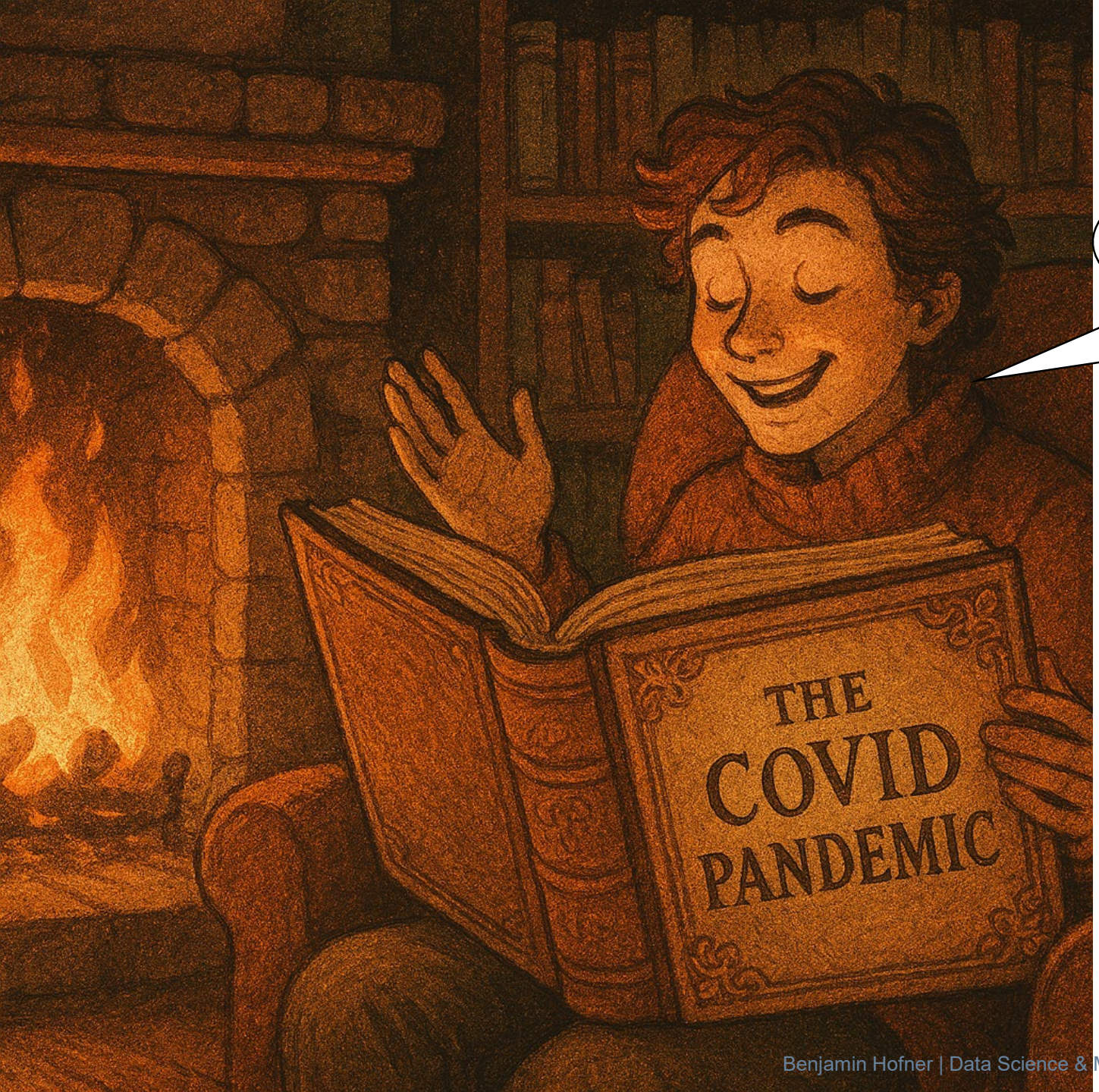
Disclaimer

Views are my own and **do not** necessarily represent the views of

- the Paul-Ehrlich-Institut (PEI),
- the European Medicines Agency (EMA), or
- any other European regulatory agency (NCA).



COVID-19 PANDEMIC



Once upon a time…

… COVID-19 emerged and quickly reached the EU

> *"The global COVID-19 pandemic arrived in Europe with its first confirmed case in Bordeaux, France, on **24 January 2020**, and subsequently spread widely across the continent. By **17 March 2020**, every country in Europe had confirmed a case."*

(Source: Wikipedia)

… COVID-19 emerged and quickly reached the EU and also Germany

> *"The global COVID-19 pandemic arrived in Europe with its first confirmed case in Bordeaux, France, on **24 January 2020**, and subsequently spread widely across the continent. By **17 March 2020**, every country in Europe had confirmed a case."*

(Source: Wikipedia)

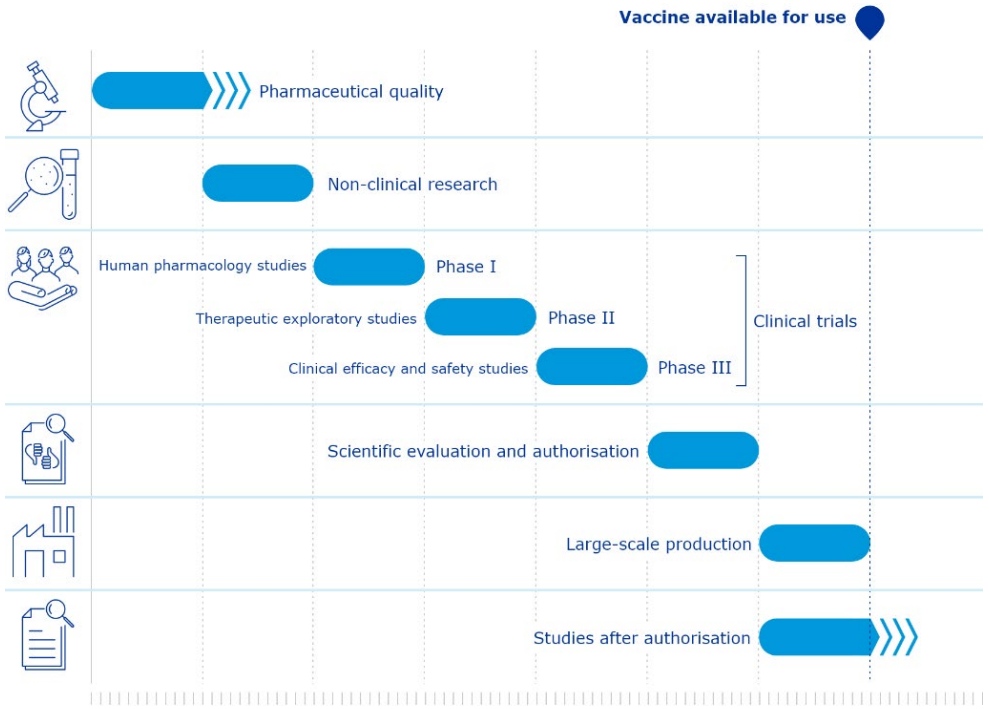
Erster Coronavirus-Fall in Deutschland | tagesthemen extra

Source: Youtube | Tagesschau (28.01.2020)

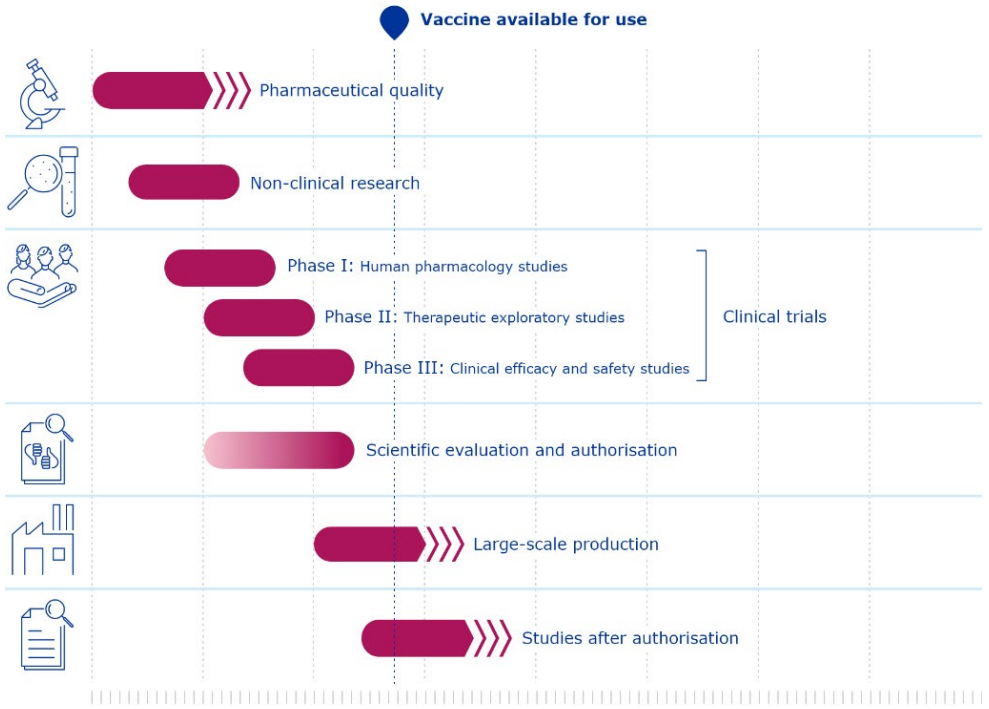
REGULATORY RESPONSE TO PANDEMIC

Vaccine development and approval

Standard process

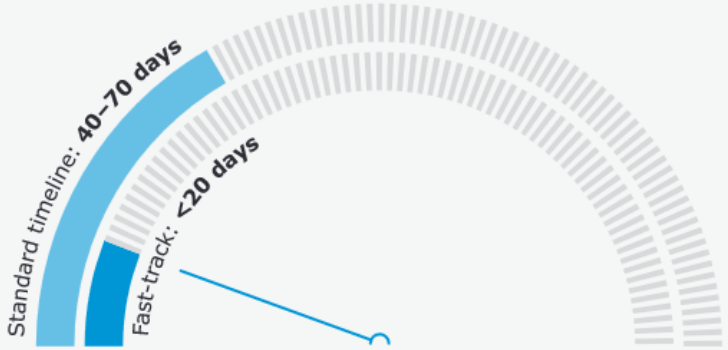


Accelerated process



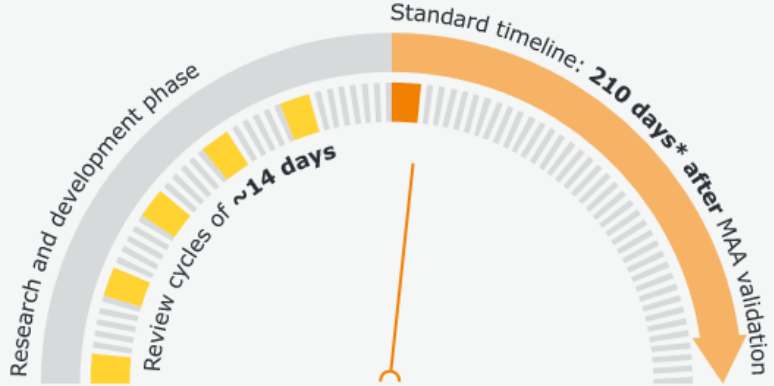
Source: EMA | COVID-19 vaccines: development, evaluation, approval and monitoring

Additional acceleration



Rapid scientific advice

EMA provides developers with prompt advice to guide on the best methods and study designs to generate the scientifically robust evidence needed to determine the safety, efficacy and quality of treatments and vaccines against COVID-19 in the shortest time possible.



Rolling review**

In a public health emergency, EMA assesses data for promising medicines or vaccines as they become available. Through rolling review, EMA can exceptionally start evaluating data while the development is still ongoing. When the medicine's development is progressed enough for a marketing authorisation application (MAA), the formal assessment procedure can take place in a very short timeframe, because the data have already been scrutinised during rolling review.

Each rolling review cycle requires around 2 weeks, depending on the amount of data

Source: https://www.ema.europa.eu/en/documents/leaflet/infographic-fast-track-procedures-treatments-vaccines-covid-19_en.pdf

Faster, but safe

MEDIZINREPORT

Ein Impfstoff, der vor COVID-19 schützt, ist essenziell, um der Coronapandemie ein Ende zu setzen. Das Paul-Ehrlich-Institut hat Maßnahmen ergriffen, um die Entwicklung und Zulassung von Impfstoffen zu beschleunigen, ohne dabei Kompromisse bei der Sicherheit einzugehen.

Dr. Ralf Wagner, Prof. Dr. Eberhard Hildt, Dr. Elena Grabski, Dr. Yuansheng Sun, Dr. Heidi Meyer, Dr. Annette Lommel, Dr. med. Brigitte Keller-Stanislawski, Dr. Jan Müller-Berghaus, Prof. Dr. Klaus Cichutek

Paul-Ehrlich-Institut, Bundesinstitut für Impfstoffe und biomedizinische Arzneimittel, Langen

→High regulatory standards were kept

Source: Deutsches Ärzteblatt | Jg. 117 | Heft 39 | 25. September 2020

LESSONS LEARNED

FOR CLINICAL TRIAL DESIGN AND REPORTING

WHAT HAS HAPPENED?

Group sequential design

	BioNTech / Pfizer	Moderna
Planned analyses (# cases)	32, 62, 92, 120, 164	53, 106, 151
Conducted analyses (# cases)	94, 170	95, 196

- Misalignment of analyses time points and additional “overrunning” at all analyses
- Possible reasons:
  - “Speed up” of pandemic due to “winter wave”?
  - Regulatory agreement?
  - Data driven choices?
    (not an issue here as all results highly significant and clinically relevant - but relevant in other cases)

Group sequential design

And what has happened?

Significant results

	BioNTech / Pfizer	Moderna
Planned analyses (# cases)	32, 62, 92, 120, 164	53, 106, 151
Conducted analyses (# cases)	94, 170	95, 196

- Misalignment of analyses time points and additional “overrunning” at all analyses
- Possible reasons:
  - “Speed up” of pandemic due to “winter wave”?
  - Regulatory agreement?
  - Data driven choices?
    (not an issue here as all results highly significant and clinically relevant - but relevant in other cases)

WHAT WAS COMMUNICATED?

Communication is key

Communication of results
(BioNTech/Pfizer)

"The first interim analysis of our global Phase 3 study provides evidence that a vaccine may effectively prevent COVID-19. This is a victory for innovation, science and a global collaborative effort," said Prof. Ugur Sahin, BioNTech co-founder and CEO. "When we embarked on this journey 10 months ago this is what we aspired to achieve. Especially today, while we are all in the midst of a second wave and many of us in lockdown, we appreciate even more how important this milestone is on our path towards ending this pandemic and for all of us to regain a sense of normality. We will continue to collect further data as the trial continues to enroll for a final analysis planned when a total of 164 confirmed COVID-19 cases have accrued. I would like to thank everyone who has contributed to make this important achievement possible."

Source: Pfizer (09. November 2020)

Communication of results
(BioNTech/Pfizer)

"The first interim analysis of our global Phase 3 study provides evidence that a vaccine may effectively prevent COVID-19. This is a victory for innovation, science and a global collaborative effort," said Prof. Ugur Sahin, BioNTech co-founder and CEO. "When we embarked on this journey 10 months ago this is what we aspired to achieve. Especially today, while we are all in the midst of a second wave and many of us in lockdown, we appreciate even more how important this milestone is on our path towards ending this pandemic and for all of us to regain a sense of normality. We will continue to collect further data as the trial continues to enroll for a final analysis planned when a total of 164 confirmed COVID-19 cases have accrued. I would like to thank everyone who has contributed to make this important achievement possible."

Source: Pfizer (09. November 2020)

"Pfizer and BioNTech today announced that, after conducting the final efficacy analysis in their ongoing Phase 3 study, their mRNA-based COVID-19 vaccine candidate, BNT162b2, met all of the study's primary efficacy endpoints. Analysis of the data indicates a vaccine efficacy rate of 95% ($p<0.0001$) in participants without prior SARS-CoV-2 infection (first primary objective) and also in participants with and without prior SARS-CoV-2 infection (second primary objective), in each case measured from 7 days after the second dose. The first primary objective analysis is based on 170 cases of COVID-19, as specified in the study protocol, of which 162 cases of COVID-19 were observed in the placebo group versus 8 cases in the BNT162b2 group. Efficacy was consistent across age, gender, race and ethnicity demographics. The observed efficacy in adults over 65 years of age was over 94%."

Souce: Pfizer (18. November 2020)

Communication of results
(Moderna)

> “The primary endpoint of the Phase 3 COVE study is based on the analysis of COVID-19 cases confirmed and adjudicated starting two weeks following the second dose of vaccine. This first interim analysis was based on 95 cases, of which 90 cases of COVID-19 were observed in the placebo group versus 5 cases observed in the mRNA-1273 group, resulting in a point estimate of vaccine efficacy of 94.5% (p <0.0001).”

Source: Moderna (16. November 2020)

Communication of results
(Moderna)

"The primary endpoint of the Phase 3 COVE study is based on the analysis of COVID-19 cases confirmed and adjudicated starting two weeks following the second dose of vaccine. This first interim analysis was based on 95 cases, of which 90 cases of COVID-19 were observed in the placebo group versus 5 cases observed in the mRNA-1273 group, resulting in a point estimate of vaccine efficacy of 94.5% (p <0.0001)."

Source: Moderna (16. November 2020)

"Moderna today announced that the primary efficacy analysis of the Phase 3 study of mRNA-1273 conducted on 196 cases confirms the high efficacy observed at the first interim analysis. The data analysis indicates a vaccine efficacy of 94.1%. (…) Vaccine efficacy has been demonstrated at the first interim analysis with a total of 95 cases based on the pre-specified success criterion on efficacy. Today's primary analysis was based on 196 cases, of which 185 cases of COVID-19 were observed in the placebo group versus 11 cases observed in the mRNA-1273 group, resulting in a point estimate of vaccine efficacy of 94.1%. "

Source: Moderna (30. November 2020)

AND THE PUBLIC REACTION?

Virologen loben vorläufige Daten

BioNTech und Pfizer sind die weltweit ersten Unternehmen, die erfolgreiche Daten aus einer zulassungsrelevanten Studie mit einem Corona-Impfstoff vorgelegt haben. Am 9. November berichteten sie über eine 90-prozentige Wirksamkeit ihres Impfstoffes auf Basis einer ersten Zwischenanalyse der entscheidenden dritten Studienphase.

Dabei handelte es sich zwar um vorläufige Daten, die Impfeffektivität kann sich also nach längerer Beobachtungszeit der Probanden noch ändern. Virologen zeigten sich dennoch begeistert, "da viele laufende Impfstudien zu Covid-19 derzeit lediglich eine Erfolgsquote von mindestens 50 Prozent voraussetzen", wie etwa der Münchner Infektologe Clemens Wendtner betonte.

Source: Tagesschau

Laut Moderna Wirksamkeit von 94,1 Prozent

Der Moderna-Impfstoff mRNA-1273 hat eine Wirksamkeit von 94,1 Prozent, wie das Unternehmen mitteilte. Das gehe aus der neuesten Analyse von Daten der klinischen Phase III-Studie hervor, an der 30.000 Menschen in den USA teilnehmen. Eine Hälfte der Teilnehmer bekommt dabei den Impfstoff, die andere Hälfte ein Placebo-Mittel. Für den vollen Impfschutz sind zwei Dosen in zeitlichem Abstand notwendig.

Insgesamt wurden unter den Probanden demnach bislang 196 Fälle der Krankheit Covid-19 bestätigt. Elf dieser Fälle entfielen auf die Gruppe mit dem Impfstoff, 185 Fälle auf die Placebogruppe. Daraus errechnet sich eine Wirksamkeit von 94,1 Prozent.

Die neuen Ergebnisse entsprechen etwa den vorläufigen Daten, die Moderna Mitte November veröffentlicht hatte. Zudem teilte Moderna mit, dass die insgesamt 30 schweren Covid-19-Verläufe in der Studie ausschließlich die Placebogruppe betrafen. Die Wirksamkeit von mRNA-1273 war laut Moderna über alle Altersgruppen hinweg ähnlich.

An Nebenwirkungen seien laut Moderna nur solche aufgetreten, die einer Grippe geähnelt hätten.

Source: Tagesschau

WHAT DO WE AS STATISTICANS SAY?

Good communication and statistical involvement is key

WHAT DO I MEAN HERE?

Issues observed with the communicated results

- **Analyses after the primary confirmatory analysis** were presented as part of the primary body of evidence for some of the vaccine trials.
- These **updated analyses** occurred within a **very short time frame** of around two weeks but were based on a substantially increased database with about twice the number of COVID-19 cases.
- **Language** in the press releases and media coverage was not well aligned, and **not suitable** for a general and wide communication

Observations

- **Language** often **is suggestive**
  - "**Final analysis**" leads to a misconception about **definitive nature** of the analysis which is usually not true

  - "**Interim analysis result**" in case of significant results at an analysis planned as interim leads to the misconception that these results are **not yet trustworthy** as they are not "final"
- **No common language** for analyses (and their value) after a "confirmatory analysis"
- Planning (and conduct) of trials could be improved by carefully **pre-specifying analyses and updates**

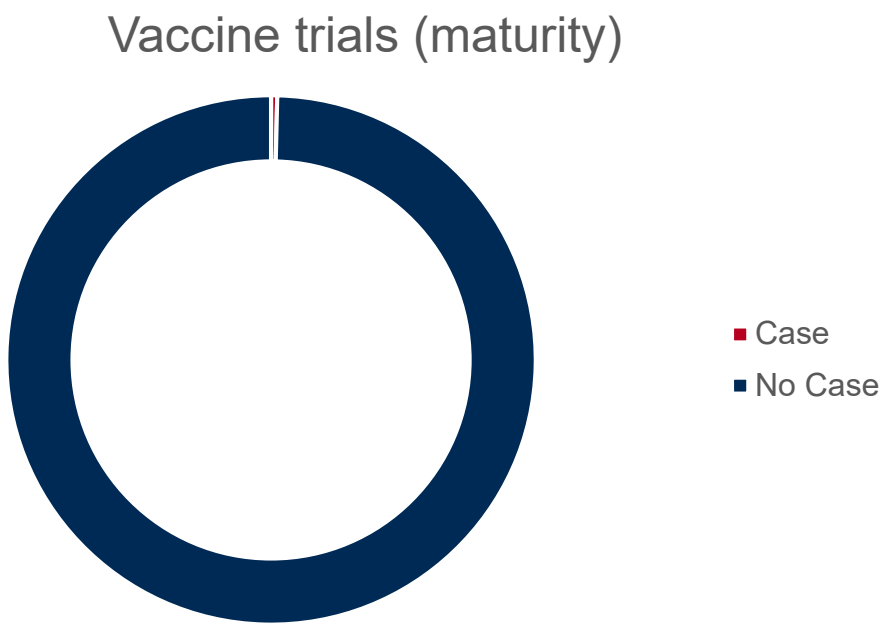
Locking for guidance
(ICH E9)

- According to the ICH E9 guideline a trial is to be **analysed** and the **primary hypothesis** is to be *"**tested when the trial is complete**" (ICH 1998)*[1] .

[1] ICH (1998). ICH E9 – Statistical principles for clinical trials. CPMP/ICH/363/96

Locking for guidance
(ICH E9)

- According to the ICH E9 guideline a trial is to be **analysed** and the **primary hypothesis** is to be *"**tested when the trial is complete**" (ICH 1998)*[1] .



About 150 cases in 30.000 to 44.000 subjects is an arbitrary choice and no natural EOT exists.

[1] ICH (1998). ICH E9 – Statistical principles for clinical trials. CPMP/ICH/363/96

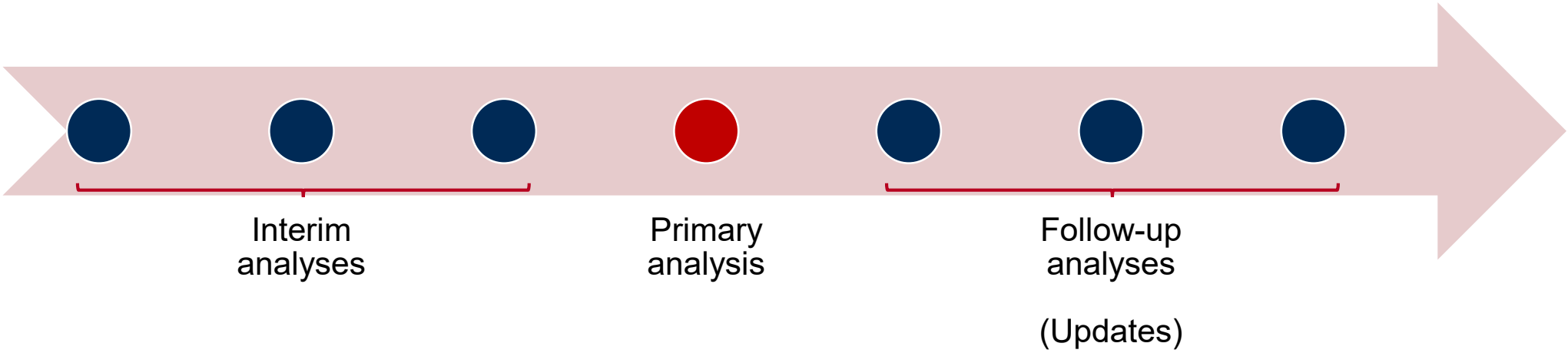
Primary analysis and end of trial

- In general, the **primary analysis** and the **end of trial** often do not coincide
  - in **group sequential designs**, which allow confirmatory testing at interim analyses (potentially without stopping the trial),
  - in **time to event trials**, where the natural end of study is almost never achieved as usually not all participants experience an event,
  - in fixed design trials with **key** (confirmatory) **endpoints at a later data cutoff**,
  - in fixed design trials, where **updated analyses or a long-term follow-up** are foreseen.

SUGGESTED TERMINOLOGY

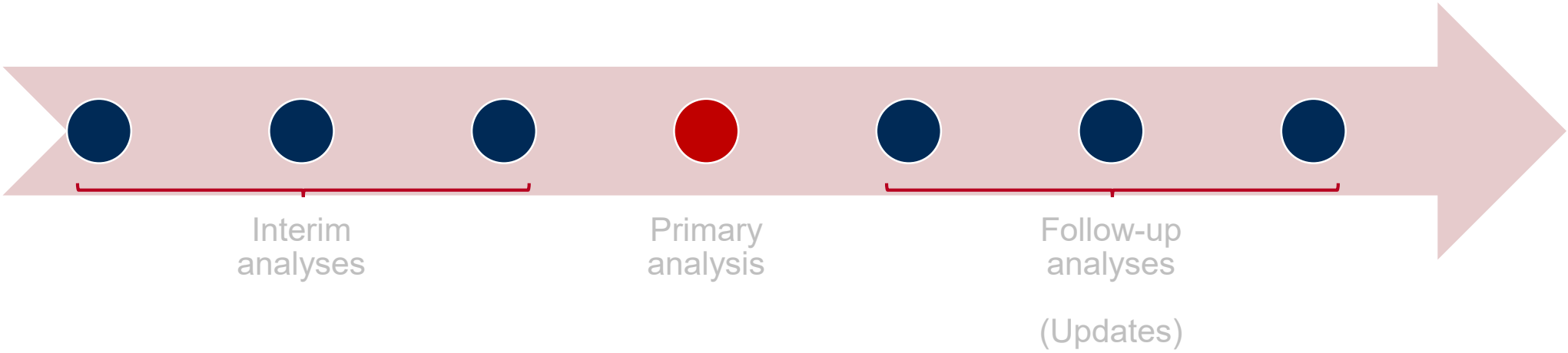
FOR GROUP SEQUENTIAL DESIGNS

Suggested terminology[1] at **planning stage**



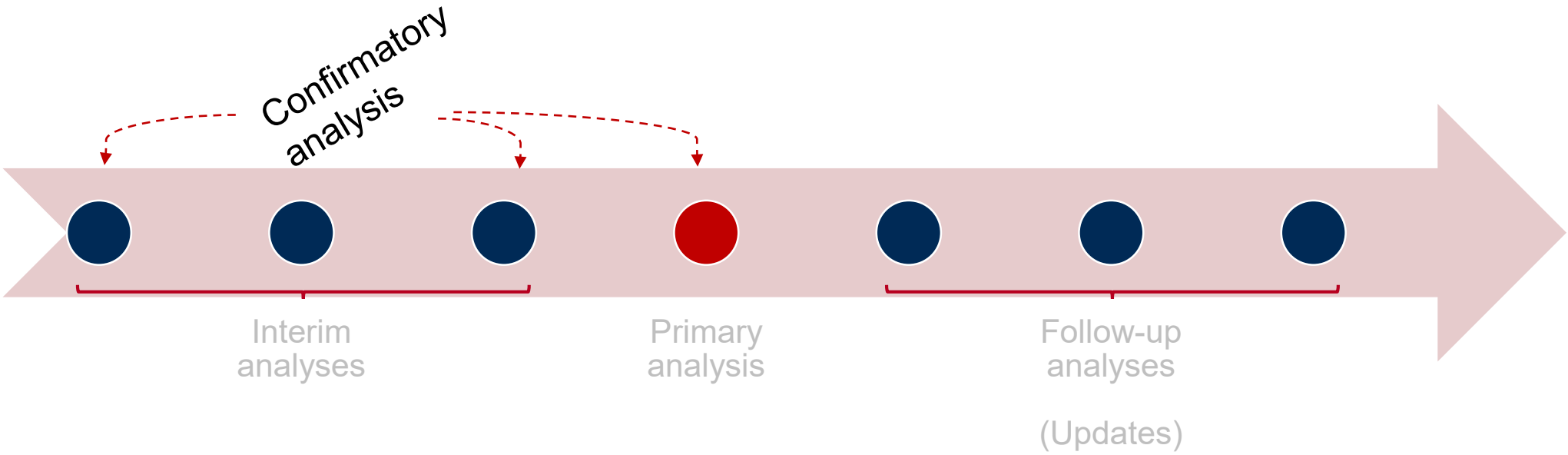
[1] per (key) endpoint

Suggested terminology[1] at **analysis stage**



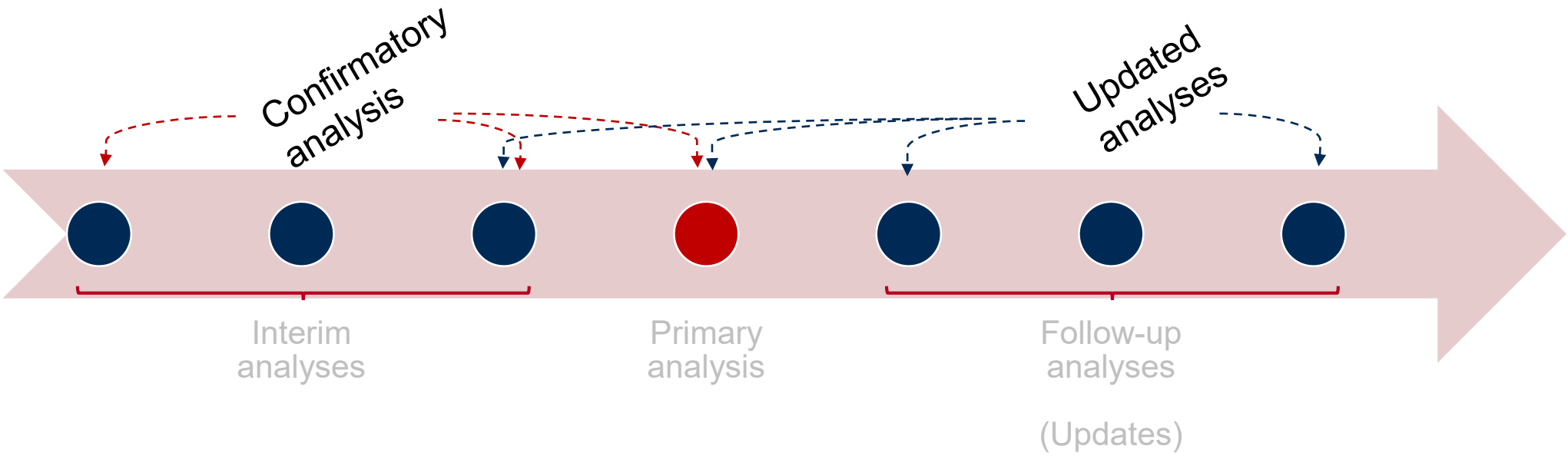
[1] per (key) endpoint

Suggested terminology[1] at **analysis stage**



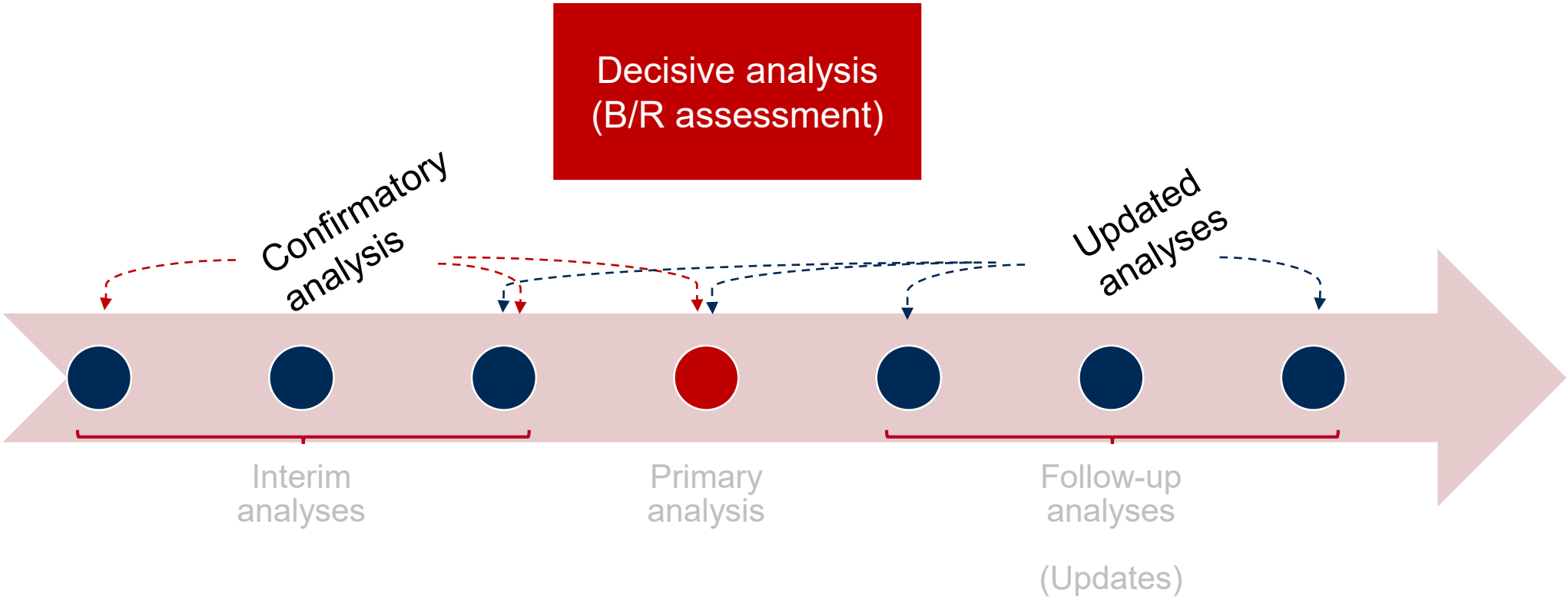
[1] per (key) endpoint

Suggested terminology[1] at **analysis stage**



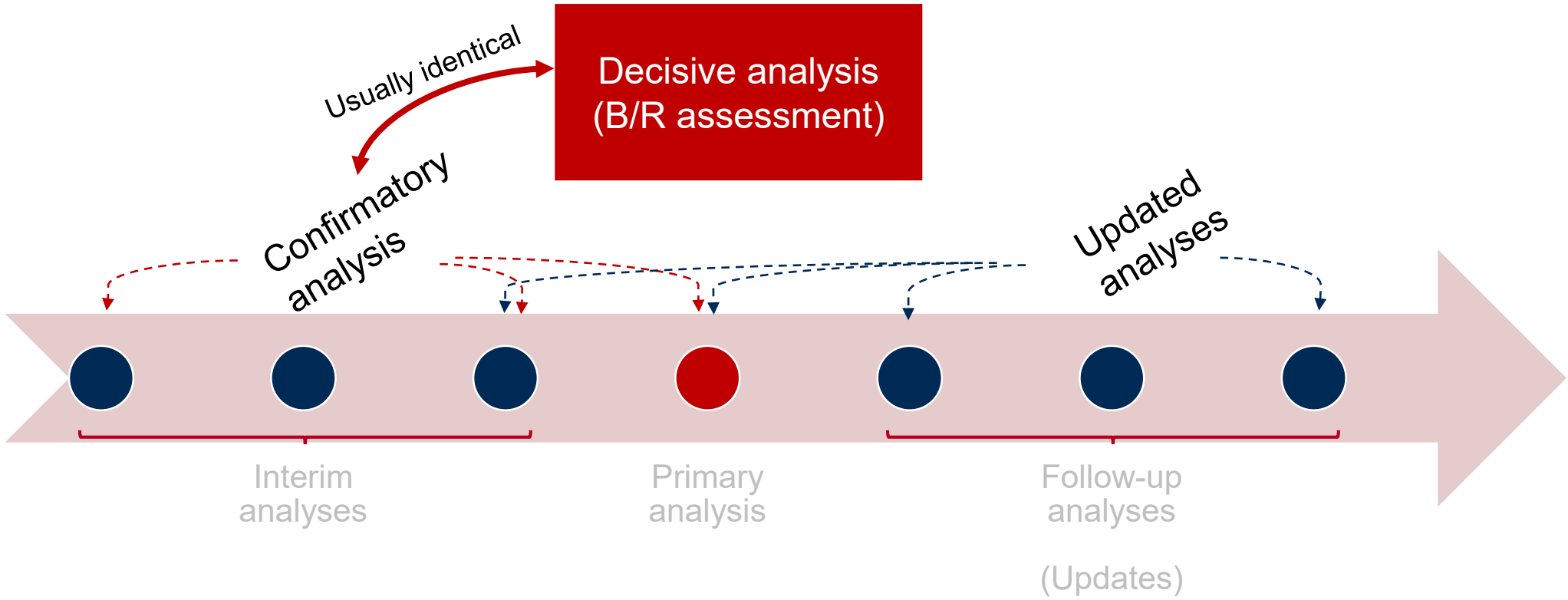
[1] per (key) endpoint

Suggested terminology[1] at **analysis stage**



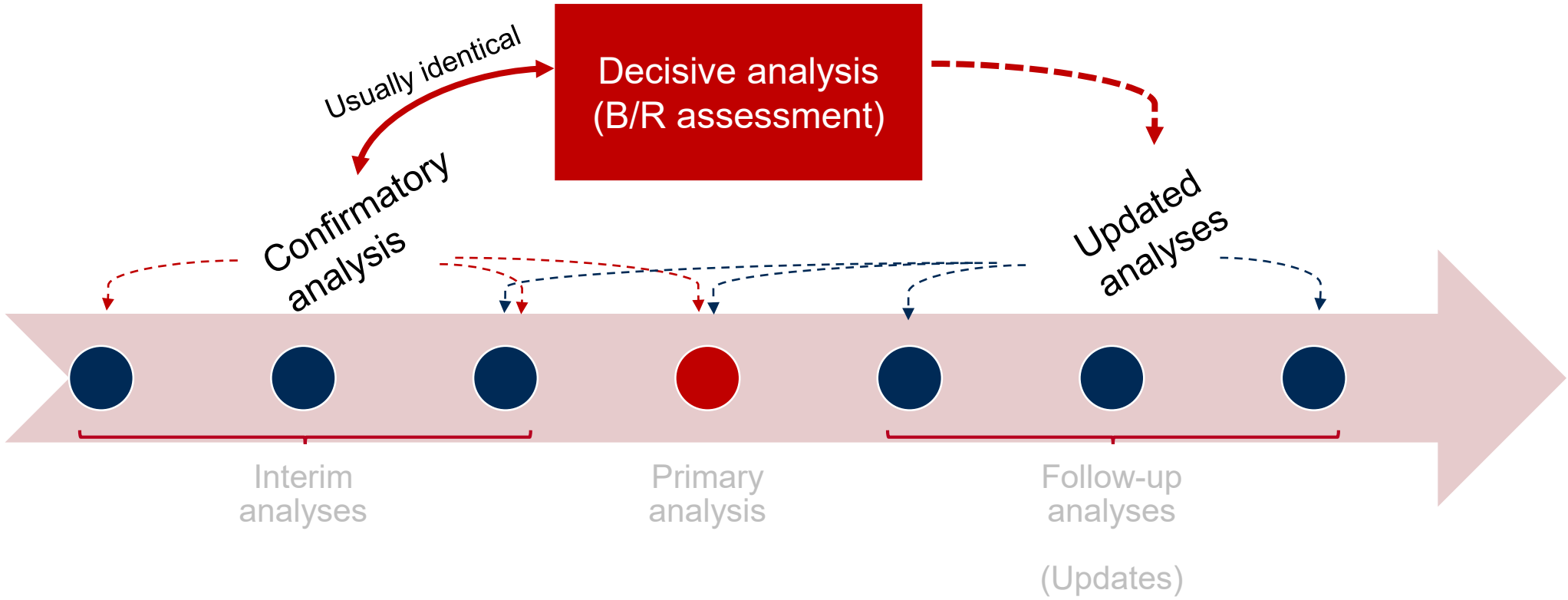
[1] per (key) endpoint

Suggested terminology[1] at **analysis stage**



[1] per (key) endpoint

Suggested terminology[1] at **analysis stage**



[1] per (key) endpoint

Updated analyses

- **Updated analyses** should collect a substantial amount of additional data and their timing and role should be thoroughly pre-planned.
- **T1E control not possible after confirmatory analysis** for that endpoint.
  - (Fully sequential approaches *might* allow to allocate significance level after confirmatory analysis)
- Report **point estimate with 95% CI**
- If a **later time point** for the same variable is of key interest (e.g. LTFU),
  T1E could be controlled e.g. in a hierarchical fashion (PEP significant then forward α to LTFU).
  - We do no longer consider this as part of the group sequential design.

Over- and underrunning

- **Overrunning** (and similarly underrunning) happens due to “random” differences of the observed from the planned information fraction. Examples include:
  - Multiple subjects were enrolled at the same time and hence reach the analysis time at the same time as well.
  - Drop-outs / withdrawals occur.
  - Multiple events happen at the same time in a time to event trial (TTE).
  - Timing of interim analysis was not adequately predicted (especially for TTE).
- Under- or overrunning *at an interim analysis* can be accounted for if a Lan-DeMets (1983) alpha-spending function is used (with the observed information fraction).
- If under- or overrunning occurs *at the primary analysis*, the remaining significance level (and only that) needs to be recomputed with the new number of events (see Wassmer & Brannath, 2016; Wolbers, Wassmer, Pahlke, online).

Pipeline information

- **Pipeline data** is due to additional data that accrues especially at interim analyses
  - Especially relevant with endpoints measured at fixed times
    (e.g. 6-minute walk test at week 26).
  - The interim is triggered with a specific information fraction
    (e.g. when 60% of subjects have reached that time point).
  - By that time more subjects were enrolled. These data for these patients is called pipeline information.
    (e.g. 80% of all planned patients were enrolled at time of interim).
- This data also includes data that accrues **during data cleaning** (which can happen only after the required number of subjects has reached the time point).
- Pipleline data **cannot be avoided** (at interim).

Overrunning and pipeline information in EMA guidelines

- EMA Reflection Paper on Adaptive Designs
  - Does not distinguish between overrunning and pipeline data.
  - “Overrunning” data needs to be taken into consideration at decision making.
  - Focus especially on estimation and clinically relevant changes in estimates.
  - Both results with and without overrunning should be presented.
  - “Overruning” data in principle overrules data from the previous (interim) analysis.

- Cave: One cannot “unreject” a hypothesis or reject it twice.
  - Approaches exist to pre-plan formal incorporation of pipeline data (e.g. Hampson and Jennison, 2013)

Key considerations for planning GSDs

- Given the many complexities, **pre-plan**…
  - the **timing** of each analysis (per endpoint and across endpoints)
  - the **role** of each analysis
- Timing should be **unambiguously and clearly defined**
  - We often see definitions which leave much room for interpretation
  - If timing is triggered by events in another endpoint it should be assured that sufficient information has accrued.
- The role of analyses should be **defined under all possible scenarios**
  - Will interim analyses after the confirmatory analysis be used for updates?
  - When will be the time point for the decisive analysis (i.e. sufficient data for benefits and risks have accrued over all subgroups and important EPs)?
- The trial should be planned such that **sufficient information for B/R** assessment is accrued **at first (potentially confirmatory) efficacy analysis**
  - Decisive analysis and confirmatory analysis usually coincide (as well as safety and efficacy analyses)
  - It should be the absolute exception to use different cutoffs for confirmation and benefit-risk assessment

Summary

- **Terminology** proposed for better communication
  - Good terminology should be used in all communications.
- **Differentiation** of overrunning, pipeline data and updates necessary
  - Over-or underrunning happen due to study procedures and nature of endpoints (“accidentally”).
  - Pipeline data is inherent to the study design and endpoint and should be considered at planning stage.
  - Updates should be planned and usually require a substantial increase of information.
- **Thoroughly plan** all analyses (i.e., interim, primary and updated analyses)
  - Specify roles of all analyses under all relevant scenarios.
  - Plan all analyses such that they accrue sufficient data for decision making (in-house as well as regulatory decision making).

When looking back…

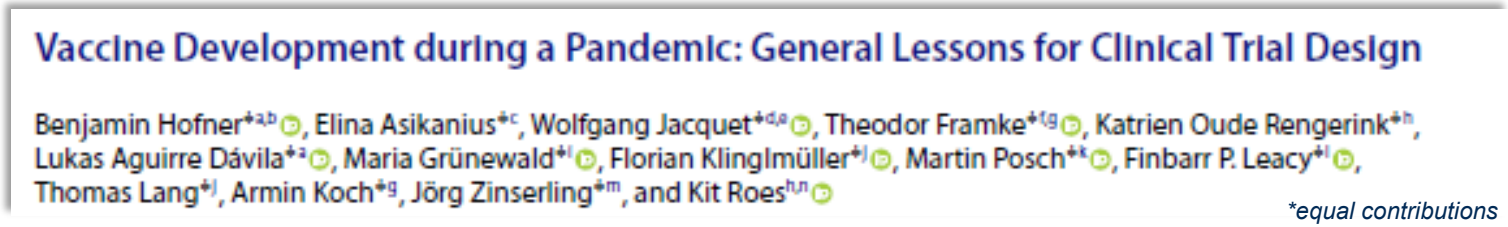
- Simple things might not be as simple as they seem.
- Communication is one of these things.
- Statisticians should not shy away from a discussion on good and scientifically sound communication.

- This colloquium is a very good example for the exchange needed **amongst statisticians** but also **with other professions.**

Congratulations to the 50th anniversary!

More on this
aka acknowledgements

This presentation is mainly based on a manuscript with colleagues from the EMA and BSWP / Methodology ESEC:

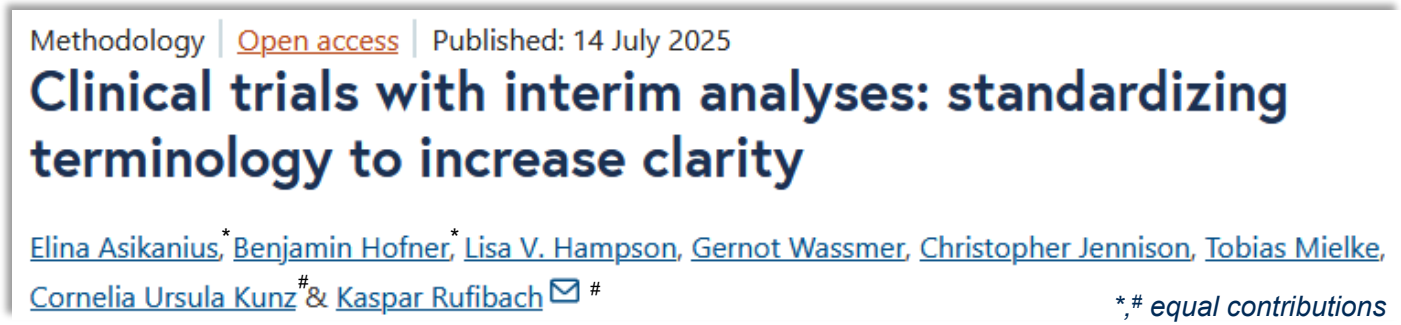

Vaccine Development during a Pandemic: General Lessons for Clinical Trial Design

Benjamin Hofner*[a,b], Elina Asikanius*[c], Wolfgang Jacquet*[d,e], Theodor Framke*[f,g], Katrien Oude Rengerink*[h], Lukas Aguirre Dávila*[a], Maria Grünewald*[i], Florian Klinglmüller*[j], Martin Posch*[k], Finbarr P. Leacy*[l], Thomas Lang*[j], Armin Koch*[g], Jörg Zinserling*[m], and Kit Roes[h,n]

*equal contributions

Statistics in Biopharmaceutical Research (2024). DOI: 10.1080/19466315.2023.2211538

And another one with colleagues from industry and academia:


Methodology | Open access | Published: 14 July 2025

Clinical trials with interim analyses: standardizing terminology to increase clarity

Elina Asikanius,* Benjamin Hofner,* Lisa V. Hampson, Gernot Wassmer, Christopher Jennison, Tobias Mielke, Cornelia Ursula Kunz# & Kaspar Rufibach # #

*,# equal contributions

Trials (2025). DOI: 10.1186/s13063-025-08942-3