Quantifying Uncertainty in the Face of Rarity: Bayesian Methods in Regulatory Contexts

Aysun Cetinyurek Yavuz

Radboud University Medical Center

Seconded expert at Dutch Medicine Evaluation Board

Co-chair of PSI Special interest group “Small populations”

Bayes Bayesian Biostatistics Leiden, NL 2025

Radboudumc

Disclaimer

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

- Radboud University Medical Center (Radboud UMC),
- Dutch Medicine Evaluation Board (CBG-MEB)
- the European Medicines Agency (EMA), or
- any other European regulatory agency.

Outline

- Small populations and challenges
- Bayesian methods in EU regulatory context
- Bayesian methods in orphan drug approvals
- Use cases
- Bayesian methods in orphan drug scientific advices
- Challenges and considerations
- Concluding remarks

Small populations

More than 30 milion people are living with a rare disease

- Rare diseases (<5 cases per 10000)
- More than half of designated orphan medicines are intended for paediatric use
- Paediatrics

75 % of the rare disases affect children,
70% have their onset during childhood

More than 6000 distinct rare diseases

230 Orphan medicines authorised in EU (2024)

9 years on average is required for a correct diagnosis

72% of rare disease are genetic origin

Challenges with small populations

- Small patient populations -> limited data
- Hard to run large RCTs
- Traditional statistical methods are often underpowered
- A lot of heterogeneity

ICH E9

Because the predominant approaches to the design and analysis of clinical trials have been based on frequentist statistical methods, the guidance largely refers to the use of frequentist methods (see Glossary) when discussing hypothesis testing and/or confidence intervals. This should not be taken to imply that other approaches are not appropriate: the use of **Bayesian** and other **approaches may be considered** when the **reasons for their use are clear** and when the **resulting conclusions are sufficiently robust.**

Guidance mentioning Bayesian methods

Acceptance likely highest for **extrapolation in pediatric** studies:

7 October 2018
EMA/189724/2018

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Reflection paper on the use of extrapolation in the development of medicines for paediatrics

Once a reduced sample size supported by extrapolation of data from a source population has been justified, this should be translated to the prospective study design through appropriate statistical approaches.
Examples could be using Bayesian methods to explicitly borrow information (e.g. from adult trials, control groups, other paediatric clinical trials).

Guidance mentioning Bayesian methods

27 August 2024
EMA/CHMP/ICH/205218/2022
Committee for Medicinal Products for Human Use

ICH E11A Guideline on pediatric extrapolation

- Different paradigms including frequentist and Bayesian may be applicable depending on the question of interest.
- Bayesian hierarchical models for synthesizing existing evidence
- If a Bayesian design that uses an informative prior is used, strict control of the type I error rate is not possible. What is an acceptable trade-off between the risks of these two errors should be considered carefully on a case-by-case basis and discussed with regulatory agencies, as appropriate.
- Dynamic borrowing using informative priors
- Effective sample size
- Need for prespecification of the model components due to complexity of Bayesian analyses and evaluate operating characteristics

ACT EU Q&A
Complex clinical trials with Bayesian design

- "In complex clinical trials, Bayesian approaches are often used for specific trial activities such as, **interim** and final analyses (including for futility and for extrapolation), **adaptations**, **pooling** of data (active, control or external), or even using **external controls**, where any such activity needs a self-standing motivation (…)"
- "Adjustment for multiple null-hypothesis significance testing (**type 1 error control**) is a central consideration in regulatory submissions for efficacy analyses (…). When using a Bayesian methodology, it is of importance that the methodology allows for an evaluation of corresponding issues, including via simulation."
- "Methodology needs to be **reasonably transparent** and its **results interpretable** (…). This is (...) why simpler analyses may be preferred over complex ones, and, for example, why external data may be more readily useful in a text discussion of a trial's context than when included in modelling."

Guidance mentioning Bayesian methods

European Medicines Agency

London, 27 July 2006
Doc. Ref. CHMP/EWP/83561/2005

Bayesian is mentioned in guideline on trials in **small populations**:

GUIDELINE ON CLINICAL TRIALS IN SMALL POPULATIONS

'Bayesian methods are a way to formally combine knowledge from previous data or prior 'beliefs' with data from a study. Such methods may be advantageous when faced with small datasets, although introducing prior beliefs is often a concern in drug regulation. A variety of reasonable prior distributions should be used to combine with data from studies to ensure that conclusions are not too heavily weighted on the prior beliefs.'

Reasons for using Bayesian methods

- Bayesian methods may be used for **pragmatic reasons** to make it possible to estimate the effect in a given situation (e.g. with small sample size) when frequentist approach is not possible or is difficult
- In all other situations they are used to **leverage prior information, e.g.** when the **prior knowledge is important** for sequential updating (e.g. probability of success)

Search algorithm

- Based on publicly available documents from **orphan drug approvals** between years of 2021- mid 2025
  - EPARs – European Public Assessment Reports (Keyword: bayes, borrow, external)
  - Only successful MAAs
  - Initial MAAs
- AI tool from EMA to search in scientific advices and by reviewing their EPARs (if available)

Limitations

- Bayesian methods are not always reported in EPARs
- Often no explicit assessment of Bayesian methods
- Methods are not described properly, details are often lacking
- Not systematic

The use of Bayesian methods in EU regulatory domain

- Bayesian methods are part of dossiers in
  - PopPK models
  - Dose finding
  - Paediatric extrapolation
  - Rare diseases – small sample issues
  - Early phase complex trials
  - Supplementary analysis -> Bayesian hirerachical models, dynamic borrowing

- Bayesian analyses were used as **primary** evidence for Comirnaty only (**not orphan):** Bayesian analyses were used as **primary** evidence: 4 interims planned, but one performed. EPAR: “Although Bayesian analysis are not usually accepted as confirmatory evidence in pivotal trials, the magnitude of the effect in this study, makes this concern redundant.”

Overview of orphan products using Bayesian methods

- Most identified use cases: **PopPK** models with **empirical Bayes** estimates are very common and covered extensively in guidance (See e.g. Guideline on reporting the results of population pharmacokinetic analyses (CHMP/EWP/185990/06))
- Bayesian analyses as **supplementary evidence**,
  - Bayesian dynamic borrowing approaches (e.g. *Bekemv*, Omjjara, Vyvgart, *Attrogy*, Kimmtrak)
  - Bayesian meta-analyses (e.g. Zilbrysq)
  - Bayesian analysis after frequentist (e.g. Nexviadyme – interpretation, *Breyanzi*, Scemblix & Ribociclib, Duvyzat, Ervebo– logistic regression with small samples)
  - Paediatric (E.g. empagliflozin – DINAMO trial)
  - Modelling of treatment activity (e.g. Vitrakvi - Additional data after CMA, Elzonris)
- **Bayesian adaptive designs**
  - futility interim analysis, (e.g. *Lumykras*)
  - Basket trial (e.g. Rozlytrek, Lumykras)

USE CASES

Bekemv (2023)

Analysis performed across trials (pooled analyses and meta-analysis)

- Paroxysmal nocturnal haemoglobinuria (PNH) is a rare, chronic, life-threatening blood disorder associated with anaemia due to haemolysis – adult patients
- Pivotal study with 42 patients, A **supplemental analysis** for the **primary endpoint** of LDH (lactate dehydrogenase) at week 27 was **pre-planned**
- **prior information from 5 historical** eculizumab **clinical studies** is used for a MAP prior for the LDH in the Soliris arm via Bayesian normal-normal hierarchical model
- The prior ESS was 23.13 subjects.
- Discounted adaptively to establish the elastic MAP (EMAP) prior.
- The particular discounting function was chosen to induce strong information borrowing when Study data are congruent to historical data (i.e. use MAP worth of 23.13 subjects as the prior) or discard the historical data altogether when a discrepancy exists, leading an adoption of a non-informative prior for mean LDH in Soliris.
- Based on prespecified criteria and LDH data at week 27, the EMAP prior for mean natural log(LDH) at week 27 for the Soliris arm was derived as normal distribution with mean 5.53 and infinite variance, a **non-informative prior** with **no borrowing** from historical data.
- Consequently, posterior week 27 GMR of ABP 959 vs. Soliris is 1.022 with respective 97.5% upper credible interval 1.205.
- **NI is concluded.**
- “Ultimately the applicant’s pre-specified algorithm resulted in **no use of historical** data sources. While the data and methodology used for this supplementary analysis leave some questions open, these are not considered critical because study is considered, in principle, to provide supportive data only.”

Lumykras (2021)

Bayesian futility interim analysis in Phase 2

- Adult patients with previously treated KRAS G12C mutated locally advanced or metastatic non-small cell lung cancer (NSCLC)
- Basket trial: An ongoing phase 1/2, open label, single-group study of sotorasib in subjects with KRAS p.G12C-mutated advanced or metastatic NSCLC, colorectal cancer, and other solid tumours.
- The interim futility analyses in a continuous manner using Bayesian predictive probability for NSCLC.
- Initial IA: after approximately 25 response-evaluable subjects
- Subsequent interim analyses after every 10 subjects
- The Go criterion was met if the probability that the true ORR exceeds the benchmark ORR is ≥ to a high probability of:
  - Go criterion for NSCLC: probability [ORR > 0.23] ≥ 80%
  - Go Criterion for CRC: probability [ORR > 0.1] ≥ 95%
- maximum planned final sample size of 105 NSCLC subjects
- Futility met if it was predicted that there is a small probability of reaching a Go Criterion upon full enrolment of 105 NSCLC given the existing observed data.
- A non-informative prior distribution of beta (1, 1) was used.
- Futility for the NSCLS was when the predictive probability of a Go decision was below 5%.
- The analysis population for the primary endpoint: Phase 2 full analysis
- Positive opinion on B/R for Conditional MAA, ongoing Phase 3 trial

Breyanzi (2022)

- Indicated for the treatment of adult patients with relapsed or refractory diffuse large B-cell lymphoma (DLBCL), primary mediastinal large B-cell lymphoma (PMBCL) and follicular lymphoma grade 3B (FL3B), after two or more lines of systemic therapy.
- Additional sensitivity analysis
- To mitigate the issue of small sample size, a Bayesian approach additionally to the classical logistic regression.
- The Bayesian approach was implemented with R MCMC package using a proper diffused prior with 0.0003 precision.
- Posterior distribution, posterior mean, median and 95% CrI of rate differences were reported.
- Trace plots for each parameter in each model were checked.
- For the classical logistic regression approach, the rate difference estimates and 95% CIs were calculated.

https://www.ema.europa.eu/en/documents/assessment-report/breyanzi-epar-public-assessment-report_en.pdf

Attrogy (2025)

most extensive description of Bayesian method in EPAR

- Indicated for the treatment of hereditary transthyretin-mediated amyloidosis in adult patients with stage 1 or 2 polyneuropathy.
- Single pivotal study with 130 patients in total
- Bayesian Augmentation: one or more of the randomised treatment arms are strengthened by combination with relevant external data that serve as an **informative** prior.
- Randomised arms without relevant, external data are augmented with a non-informative prior.
- The registry data were compared vs the pivotal trial data and used to augment the trial data in a Bayesian analysis.
- Bayesian augmentation of the study data with the diflunisal registry data served to enhance treatment effect estimates versus placebo; the probability diflunisal therapy is superior to placebo after 24 months was 99.6% and 97.2% for **Kumamoto score** and **mBMI** respectively
- The applicant: The registry data are highly consistent with the results of the pivotal clinical trial and provide valuable reassurance regarding their reliability.
- The assessment: Given the differences in study population, this was not considered adequate. It is noted that the overall registry population included mostly elderly patients. Therefore, the overall value of the Registry is considered very limited.

https://www.ema.europa.eu/en/documents/assessment-report/attrogy-epar-public-assessment-report_en.pdf

Bayesian methods in scientific advices

- In scientific advices there are more proposals with (compared to MAA)
  - Bayesian designs with interim analysis
  - Small sample issues (e.g. Bayesian logistic regression)
  - Bayesian dynamic borrowing of external controls using power priors, rMAP or Bayesian hierarchical models
  - Bayesian dynamic borrowing for extrapolation to paediatrics

- Adaptive Bayesian design (e.g. **Letter of support** for the Global Platform Study of Novel Medicines in Paediatric and Adolescent Relapsed and Refractory B-cell Non-Hodgkin Lymphoma (**Glo-BNHL platform**))

https://www.ema.europa.eu/en/documents/other/letter-support-global-platform-study-novel-medicines-paediatric-and-adolescent-relapsed-and-refractory-b-cell-non-hodgkin-lymphoma-glo-bnhl-platform_en.pdf

Some issues identified from the EPARs and SA

- No prespecification of supplementary analysis (frequentist not significant, then additional analysis using Bayesian methods)
- No prespecification of priors especially when borrowing is planned to be included
- Higher type I error rate with informative priors

- In almost all cases the applicant was advised to use frequentist procedure for primary analysis

➢Bayesian designs for **primary (efficacy) analyses and confirmatory evidence** not (really) accepted.

Challanges and considerations

- More heterogeneity in small populations
- Complexity of the model specification => weights in priors, mean and SD of the noninformative component
- Pre-specification of all components and/or the process
- Data quality when external data is planned to be used
- Prior selection
- Computational complexity
- Good **communication and** interpretations of results
- Acceptance and standardization in regulatory frameworks

Acceptability of Bayesian methods

- Bayesian methodology in clinical trials change the current paradigm
- Self-standing evidence is replaced with complex combination of data with prior information, concern especially for **informative priors**.
- **Justification** of the models, their suitability and frequentist properties (operating characteristics) are needed
- Extent and acceptability should likely be judged on a case-by-case basis
- Early interaction with the regulators is recommended

➢ More detailed guidance on acceptance and requirements of Bayesian methods in regulatory assessment is upcoming (EMA reflection paper on Bayesian methods)

Concluding remarks

- Bayesian statistics have a place in regulatory assessment and decision making for drugs intended for “small populations”
- Their use in pivotal evidence generation requires more rigour especially in the design stage