EFSPI EUROPEAN FEDERATION OF STATISTICIANS IN THE PHARMACEUTICAL INDUSTRY
Representing Statistical Associations in Europe

Evidence Synthesis for HTA

Squaring the Circle: bridging innovation with application

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

EFSPI EUROPEAN FEDERATION OF STATISTICIANS IN THE PHARMACEUTICAL INDUSTRY
Representing Statistical Associations in Europe

Indirect treatment comparisons

Is there a gap in the market?

Lytske Bakker, Open Health

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

Content

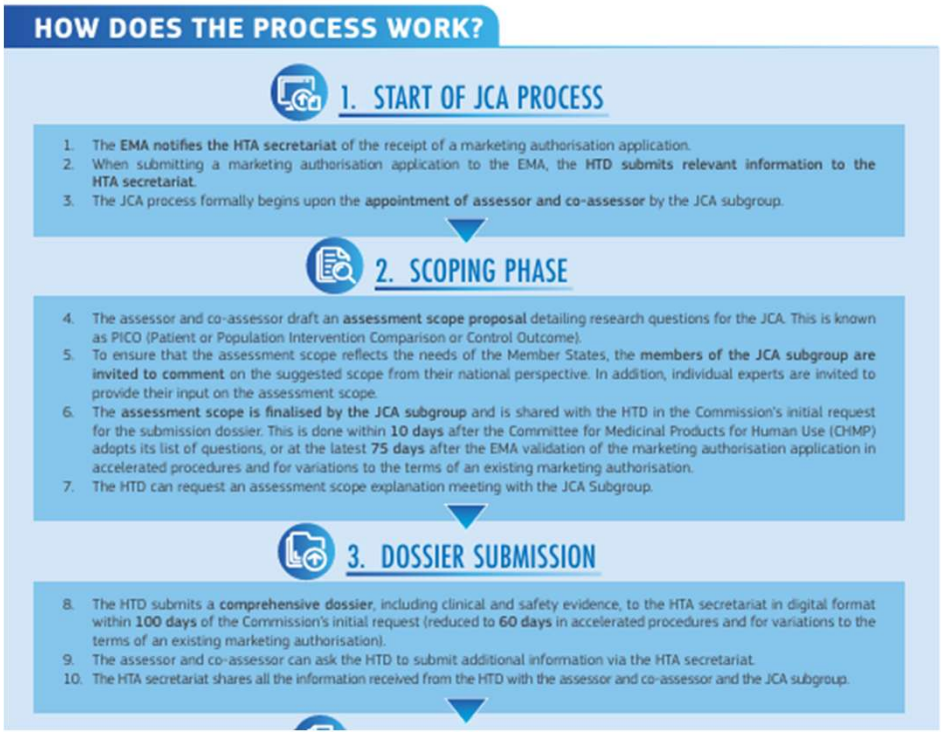
- Indirect Treatment Comparisons- the methods
- Review of indirect treatment comparisons
    - An overview of methods used
    - Criticism of methods adopted
- Introduction of the panel

Acknowledgements

The content of this presentation was developed together with Bodille Blomaard, Svenja Petersohn and Claire Ainsworth

Background

- The Joint Clinical Assessment (JCA) gives companies 100 days to submit a *comprehensive* dossier
- Many Patient, Intervention, Comparator, Outcome (PICOs) combinations?
  - Many indirect treatment comparisons (ITC) in 100 days
- Swift adoption and use of best practice methods is essential to meet with JCA requirements



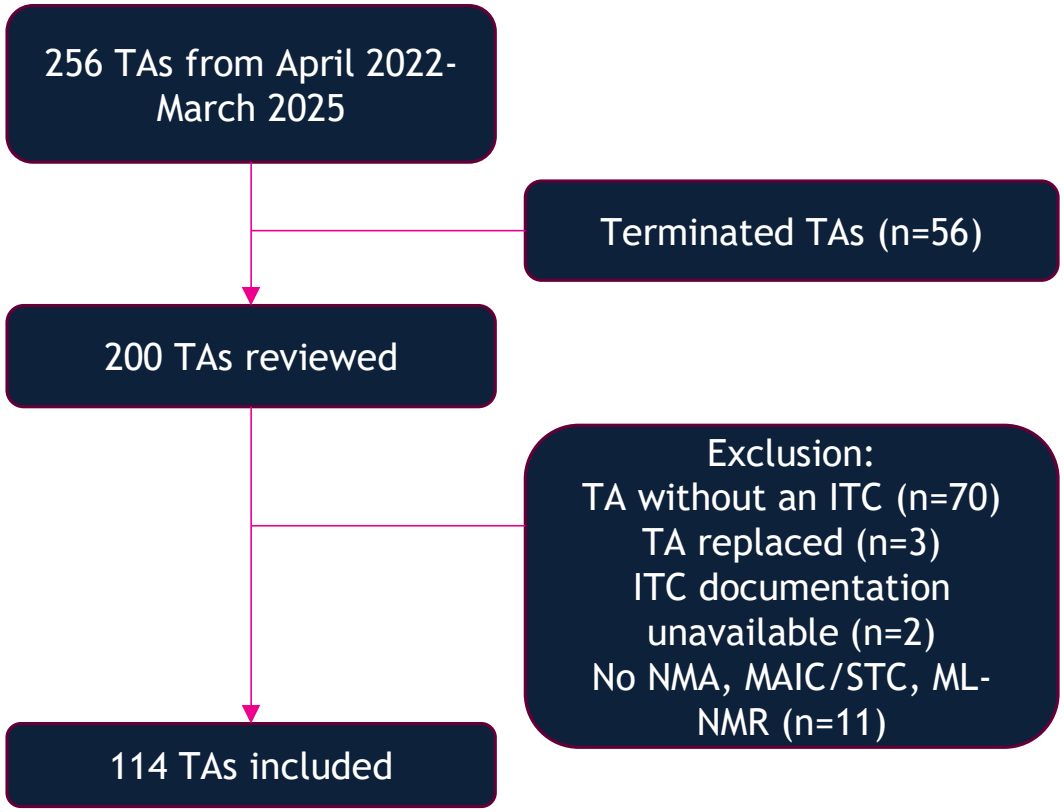
ITCs – methods

	NMA	MAIC	STC	ML-NMR
Comparison	Multiple (n ≥ 2) treatment effects	Single comparison	Single comparison	Multiple (n ≥ 2) treatment effects
Data needs	Aggregate only	IPD (1 study)	IPD (1 study)	IPD (≥1 study)
Treatment effect can be estimated for	Assumed population of included studies	Population of aggregate data	Population of aggregate data	All trial populations included in the analysis, **and** within populations where treatments have not been observed*
Adjustment for population differences to reduce bias in estimates	✓	✓	✓	✓
Feasible with single-arm trials	✗	✓	✓	✗
Can be performed without a common comparator arm	✗	✓	✓	✗
Acceptability	"Gold standard"	Most frequently used PA-ITC	Recommended for unanchored ITC	Newest methodology, growing traction

* Providing relevant covariate information for that population is available

ITC methodologies- the status quo

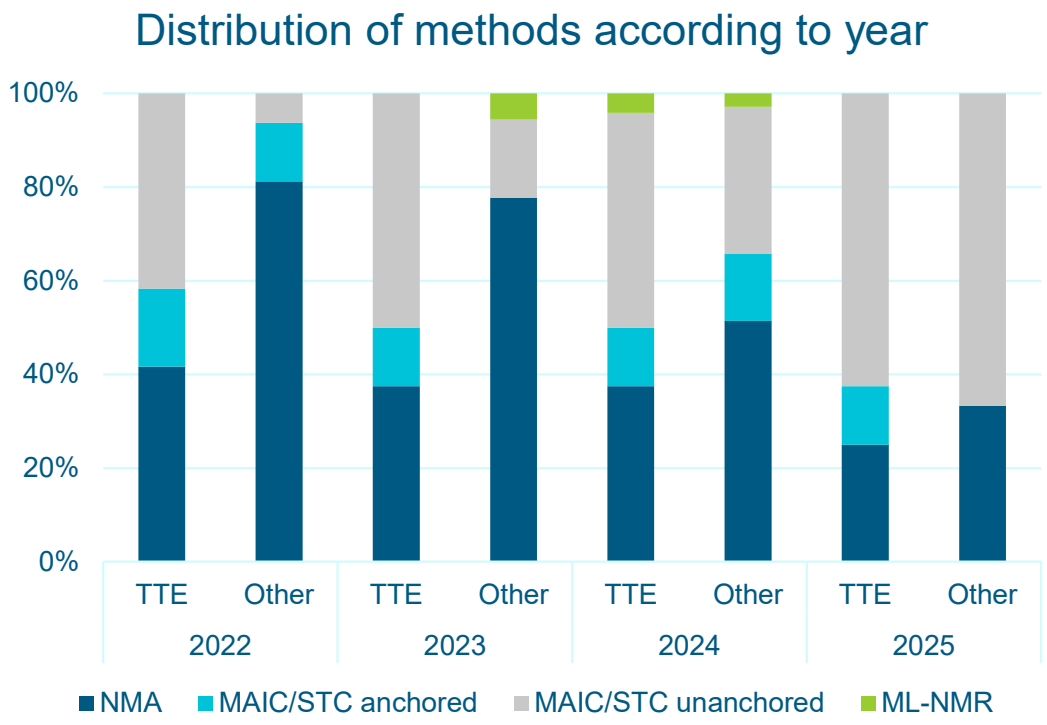
- Identify methods adopted and criticism expressed by the National Institute for Health and Care Excellence (NICE)
- Submissions from April 2022-March 2025
- Exclusion criteria:
  - Termination of the appraisal
  - No ITC performed
- Including NMAs/MAICs/STCs/ ML-NMRs



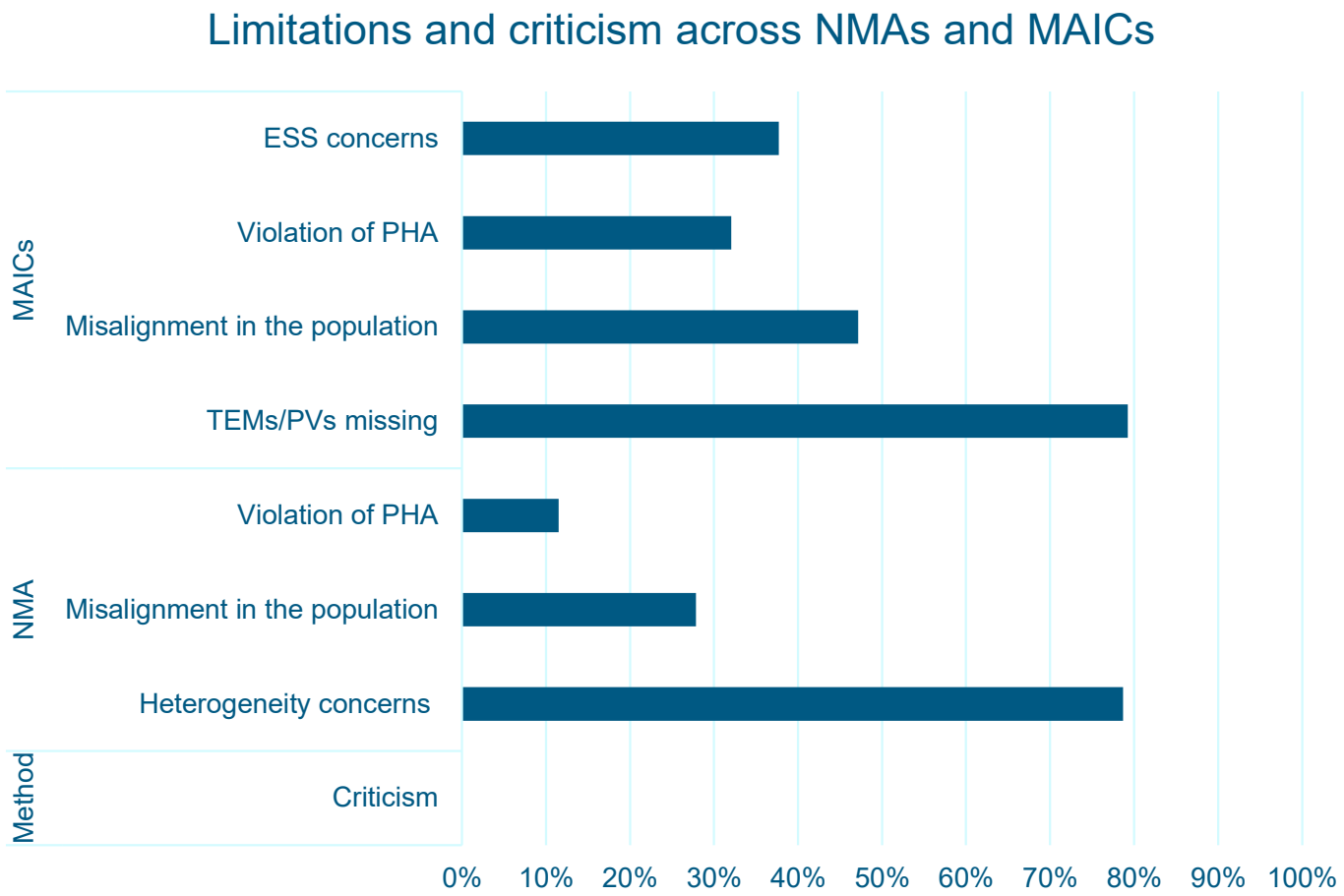
ITCs- Current practice

- NMAs were most often used followed by unanchored MAICs
- STCs were rarely used
- ML-NMRs were only recently available and thus seen less often

- The outcomes reported were most often binary (n=56) or time-to-event (TTE) data (n=60)



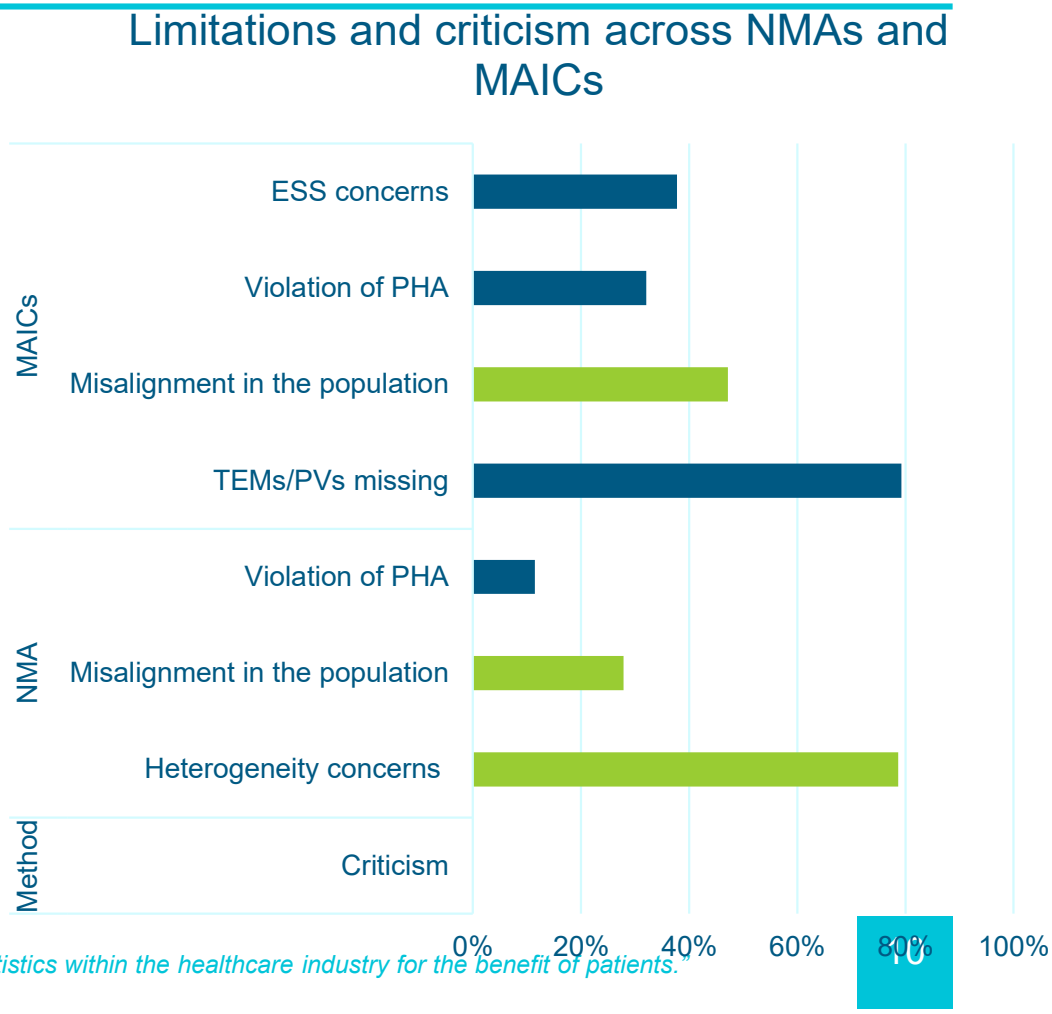
ITCs- Criticism and challenges



ML-NMR: A need for better methods?

- NMAs and MAICs most often used
- Challenges identified:
  - Misalignment between populations (47% MAICs and 28% NMAs)
  - Both MAICs and NMAs in 13% of TAs
  - Heterogeneity in NMAs in 79% of TAs

- **Nicky Welton** will introduce ML-NMRs as a method and discuss the underlying assumptions



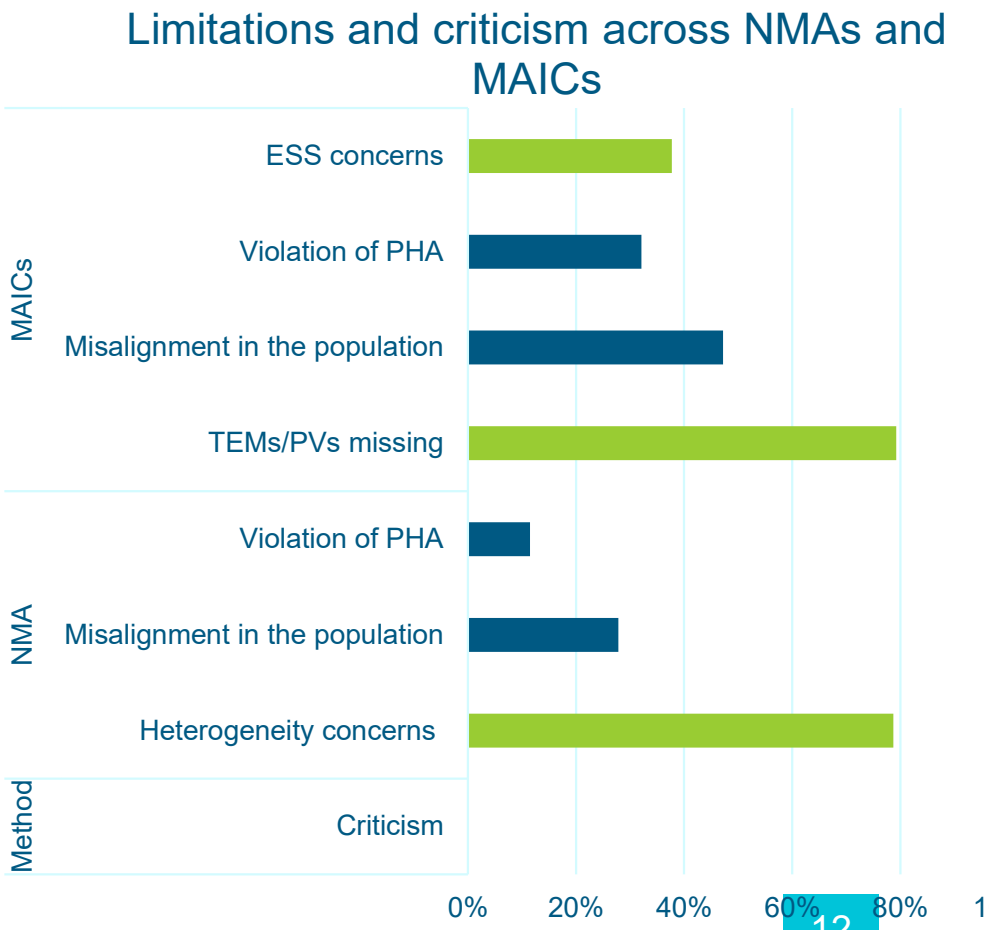
ML-NMR: How to ensure adoption?

- ML-NMRs coincide with an increased layer of complexity
  - STC's (n= 6) are rarely used despite recommendations[1,2]
  - Companies seem to avoid random effects in 23% of submissions
  - Proportional hazard assumption is often criticized (32% of MAICs)

- **Min-Hua Jen** will discuss challenges when considering the use of ML-NMR from an industry perspective

Better methods? Better planning!

- Misalignment between planning of trials and evidence required for HTA dossiers?
  - Missing treatment effect modifiers and prognostic variables in 79% of MAICs
  - Effective sample size concerns in 38% MAICs
  - 74% of MAICs is unanchored
- **Gregory Chen and Anders Gorst-Rasmussen** will discuss quantitative scenario planning for ITCs



Better methods *or* better planning

- Efficiency is key
  - JCA requires companies submit their dossier within 100 days after receiving their PICO's
  - Not *another* sensitivity analysis; 33% of ITCs presented multiple methods

- Advance HTA decision-making
- **Keith Abrams** will conclude with a discussion on how novel methods and better planning can result in better evidence dossiers

Speakers and topics

- **Nicky Welton-** Multilevel network meta-regression- the method
- **Min-Hua Jen-** Multilevel network meta-regression- the new golden standard for industry?
- **Gregory Chen and Anders Gorst-Rasmussen-** Adaptive integration and software engineering to drive planning and execution of ITCs for EU HTA- The early bird gets the worm
- **Keith Abrams-** Better methods or better planning? - Improving HTA decision-making
- Panel discussion

EFSPI EUROPEAN FEDERATION OF STATISTICIANS IN THE PHARMACEUTICAL INDUSTRY
Representing Statistical Associations in Europe

Multilevel network meta-regression

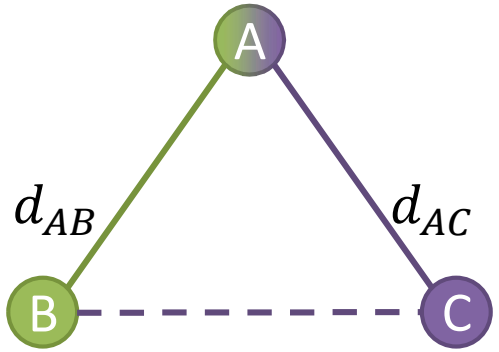
The method

Nicky Welton, University of Bristol

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

Indirect Comparisons: Assumption

- Assumes constancy of relative effects:
  - $d_{AB(AB)} = d_{AB(AC)}$
- Biased if there are imbalances in effect modifiers between AB and AC
  - $d_{AB(AB)} \neq d_{AB(AC)}$
- "Anchored" indirect comparisons are robust to differences in purely prognostic factors between trials
  - due to randomisation

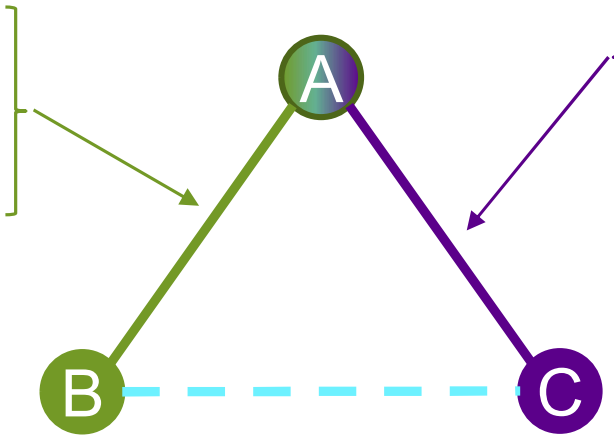


Population Adjusted Indirect Comparisons

- Seek to adjust for imbalance in effect modifiers (EM)s
- Create a fair comparison in a specific target population
  - ie for a given set of EMs
- Common scenario: limited IPD

AB trial: IPD

Y_i	T_i	X_{1i}	X_{2i}	...



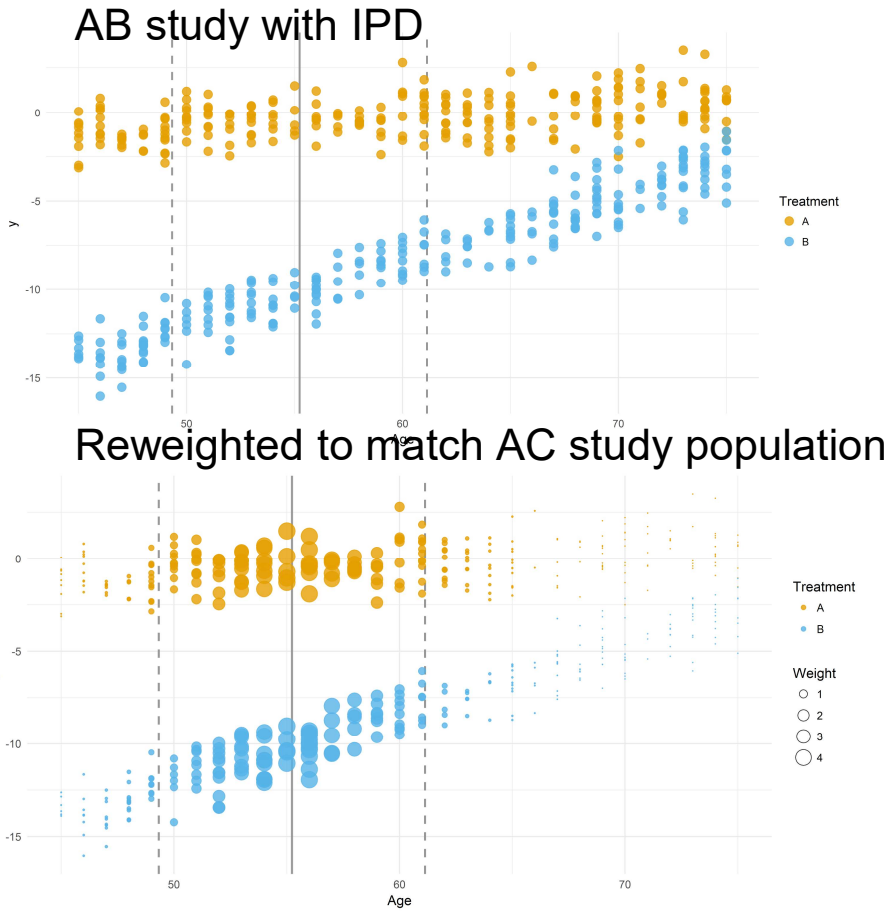
AC trial: aggregate data for outcomes and covariates

$\bar{Y}_A, \bar{Y}_C, \bar{X}_1, \bar{X}_2, \ldots$

$\sigma_A, \sigma_C, f_X(\cdot)$

Matching-Adjusted Indirect Comparison (MAIC) Signorovitch et al. (Pharmacoeconomics 2010)

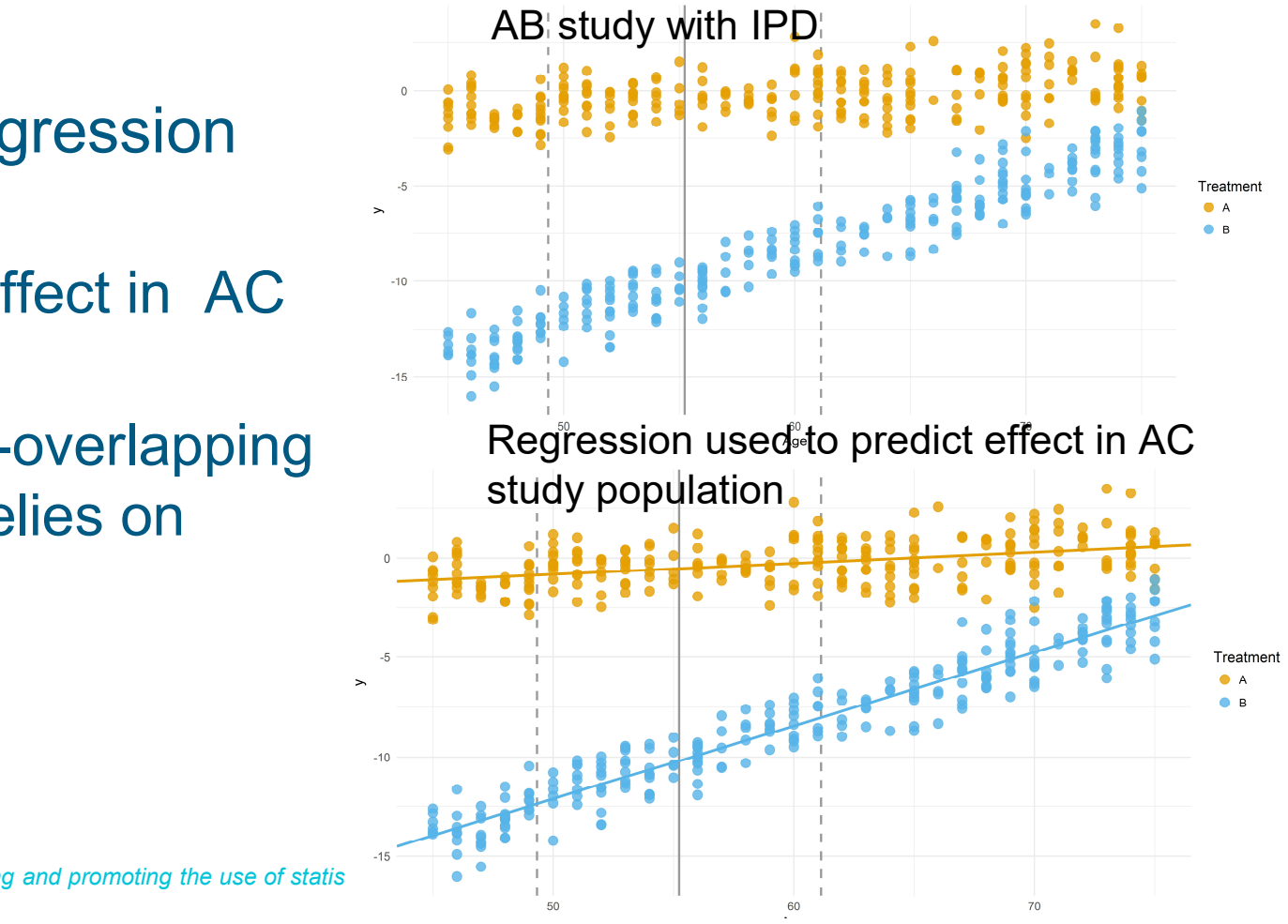
- Population reweighting method
  - similar to propensity score weighting
- Effective sample size depends on overlap between studies
  - can give unstable / imprecise estimates
- Frequently used in NICE TAs



"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

Simulated Treatment Comparison (STC) Ishak et al. (Pharmacoeconomics 2015)

- Create an outcome regression model in the AB trial
  - predict AB relative effect in AC population
  - can be used for non-overlapping populations … but relies on extrapolations …



Limitations of MAIC and STC

- Both produce estimates of relative treatment effect that are specific to the AC population
  - May not represent the target population for the decision
- MAIC performs best when sufficient overlap in trial pops
  - … but then population adjustment isn't needed!
- Neither generalise to larger networks
  - Performing several MAICs/STCs from one IPD study into several aggregate studies is invalid
    - Mixing different populations, different adjustments, multiple uses of the same data etc.

Multilevel Network Meta-Regression (ML-NMR) Phillippo et al (JRSSA 2020)

- Define an individual-level regression model

individual i, study j, treatment k

Individual:

$$y_{ijk} \sim \pi_{\text{Ind}}(\theta_{ijk})$$

"link" eg logit

$$g(\theta_{ijk}) = \eta_{jk}(\boldsymbol{x}_{ijk}) = \mu_j + \boldsymbol{x}_{ijk}^{\mathsf{T}}(\boldsymbol{\beta}_1 + \boldsymbol{\beta}_{2,k}) + \gamma_k$$

covariates; covariate main effect; covariate x treatment interaction; treatment main effect

- Aggregate-level model obtained by integration (Quasi MC)

Aggregate:

$$y_{\bullet jk} \sim \pi_{\text{Agg}}(\theta_{\bullet jk})$$

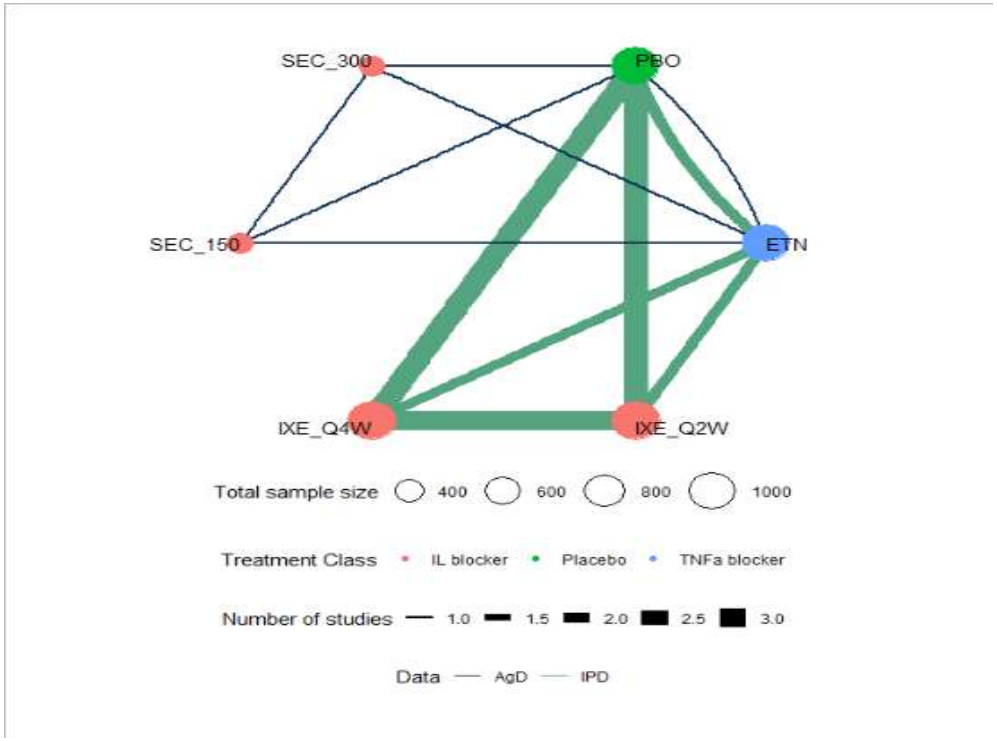
$$\theta_{\bullet jk} = \int_{\mathfrak{x}} g^{-1}\left(\eta_{jk}(\boldsymbol{x})\right) f_{jk}(\boldsymbol{x})\, d\boldsymbol{x}$$

covariate distribution

  - Use copula for covariate distribution, and Quasi MC integration

Multilevel Network Meta-Regression (ML-NMR)

- General framework with special cases:
  - IPD network meta-regression with full IPD
  - Standard NMA with no adjustment
- Can be used in networks of all sizes
- Implemented in `multinma` R package



Relative Effects in Target Decision Population Phillippo et al (MDM 2023)

- Integrate relative effects over covariate distribution to obtain relative effects in target population

$$\int_{X} \left(x^T \beta_{2,k} + \gamma_k \right) f(x) dx$$

- The target population could be represented by
  - A registry data set
  - An observational study
  - An RCT

ML-NMR: Identifiability

- Independent EM interactions
  - Requires IPD, or several Aggregate data studies at different covariate values, on each treatment
- Exchangeable EM interactions
  - Similar data requirements to independent EM interactions, plus multiple treatments needed to estimate shared variance term
- Common/shared EM interactions
  - May be justified for treatment classes

ML-NMR Assumptions

- No unobserved effect modifiers (EMs)
  - Assumed by all population adjustment methods
- Shared EM assumption
  - for identifiability in small networks
  - MAIC and STC also assume this for AC estimates to generalise to other populations
- Form of joint covariate distribution in the AgD
  - Same marginal form and correlation structure as IPD
- Valid to extrapolate outside range of IPD
  - also assumed by STC (MAIC cannot extrapolate)

ML-NMR: Recent Developments

- ML-NMR has been developed for time-to-event outcomes
  - proportional and non-proportional hazards models
  - parametric and flexible (M-spline) survival models
  - implemented in latest version of `multinma`
  - https://arxiv.org/abs/2401.12640
- A beta-version of `multinma` allows the user to enter summary regression coefficients when IPD is not available
- Using ML-NMR with disconnected networks and single arm trials (unanchored indirect comparisons)
  - Work-in-progress (Perren ISCB abstract 2024)

EFSPI EUROPEAN FEDERATION OF STATISTICIANS IN THE PHARMACEUTICAL INDUSTRY
Representing Statistical Associations in Europe

Multilevel Network Meta-Regression (ML-NMR)

The New Golden Standard for Industry?

Min-Hua Jen, Manoj Khanal, Michael David Sonksen, Eli Lilly

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

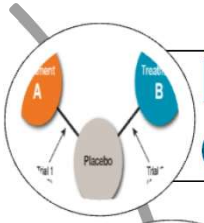
Disclaimer

The content presented is intended for informational purposes only and reflects the views of the presenter based on available data and research

Outline

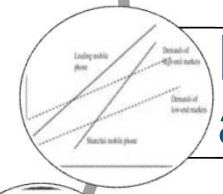
- Growing Use of ML-NMR in Industry
- Challenges in Specifying Treatment Interaction Effects in RCTs and ML-NMR
- Implementation Challenge in Practice
- ML-NMR application in time to event endpoints
- Final thoughts

Growing Use of ML-NMR in Industry

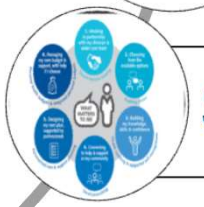


Increasingly applied in population-adjusted indirect treatment comparisons (ITCs)

Especially valuable in large networks with multiple comparators



Enables adjustment for effect modifiers and prognostic factors across studies



Supports personalized and population-level decision making

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

Challenges in Specifying Treatment Interaction Effects in RCTs and ML-NMR

- Difficulty Identifying Treatment Interactions (Harrell, 2020*)
  - RCTs powered for average effects, not interactions
  - Even balanced covariates → interaction effects 4× less precise
  - Power effectively 16× lower for differential effects (Gelman)
  - Multiple testing risks false positives
  - Requires a clear, pre-specified rationale

- Subgroup Analysis Limitations
  - Does not prove heterogeneity of treatment effect
  - Often yields unreliable estimates

Harrell 2020 https://www.fharrell.com/post/varyor/#difficulty-of-the-task

Challenges in Specifying Treatment Interaction Effects in RCTs and ML-NMR

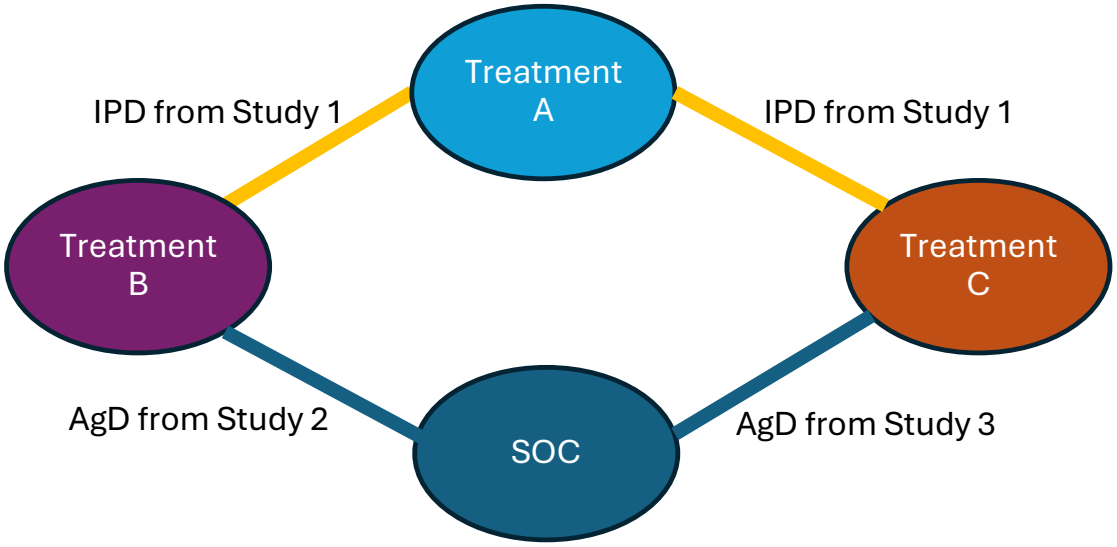
- ML-NMR Assumptions (1 RCT IPD)
  - Interaction is real (not data-mined)
  - Shared effect modifiers
  - Ideally need individual patient data (IPD) from each treatment class

- Implications
  - May limit ML-NMR use vs. standard NMA
  - Still useful as sensitivity analysis

Implementation Challenge in Practice

- Requires advanced statistical expertise in Bayesian hierarchical modelling
- Limited transparency due to proprietary IPD use
- High computational burden and convergence issues
- Performance hinges on covariate selection and model specification

ML-NMR application in time to event endpoints



ML-NMR Model	Random or Fixed	DIC
Piecewise exponential	random	2899.77
Log-logistic	random	2901.47
Log-logistic	fixed	2903.09
Piecewise exponential	fixed	2907.88
Gamma	fixed	2908.77
Mspline with 2 degrees	random	2910.7
Log-normal	random	2911.63
Log-normal	fixed	2911.66
Mspline with 3 degrees	random	2911.93
Weibull	random	2912.22
Mspline with 2 degrees	fixed	2919.29
Mspline with 3 degrees	fixed	2919.65
Weibull	fixed	2920.65
Mspline with 1 degree	random	2925.23
Gompertz	random	2933.67
Mspline with 1 degree	fixed	2934.48
Gompertz	fixed	2941.26
Exponential	random	2971.49
Exponential	fixed	2973.47

Challenges in Applied Settings

- Trade-offs between model performance and computation speed
  - Higher number of integration points means better fit but much slower speed
- Algorithms can be slow even with small networks
  - We experienced 8+ hours for 300k samples in a 3-study network
  - Running all combinations of Fixed/Random effects, different covariates and different likelihoods leads to more computationally intensive tasks
- High memory usage
  - Even in simple networks, posterior samples and predictions may use 80+ GB
  - This becomes even more challenging when using a computing cluster
- Convergence
  - Some models may never converge or require advanced knowledge in stan

Final thoughts

- **ML-NMR is not yet the golden standard**, but a **strong candidate** for HTA-relevant comparisons
- Requires:
  - Data availability
  - Technical capacity
  - Cross-functional collaboration
- Statisticians play a central role in making it accessible and acceptable to HTA bodies
- Balance innovation with transparency and interpretability

EFSPI EUROPEAN FEDERATION OF STATISTICIANS IN THE PHARMACEUTICAL INDUSTRY
Representing Statistical Associations in Europe

Early integration and software engineering to drive planning and execution of ITCs for EU HTA

The early bird gets the worm

Gregory Chen, MSD Switzerland
Anders Gorst-Rasmussen, Novo Nordisk A/S

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

Disclaimer

- Viewpoints are our own and not necessarily those of our companies
- All illustrations are AI-generated

Getting the foundation right

“We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients.”

What if we don’t get the foundation right?

1. Mismatched Evidence

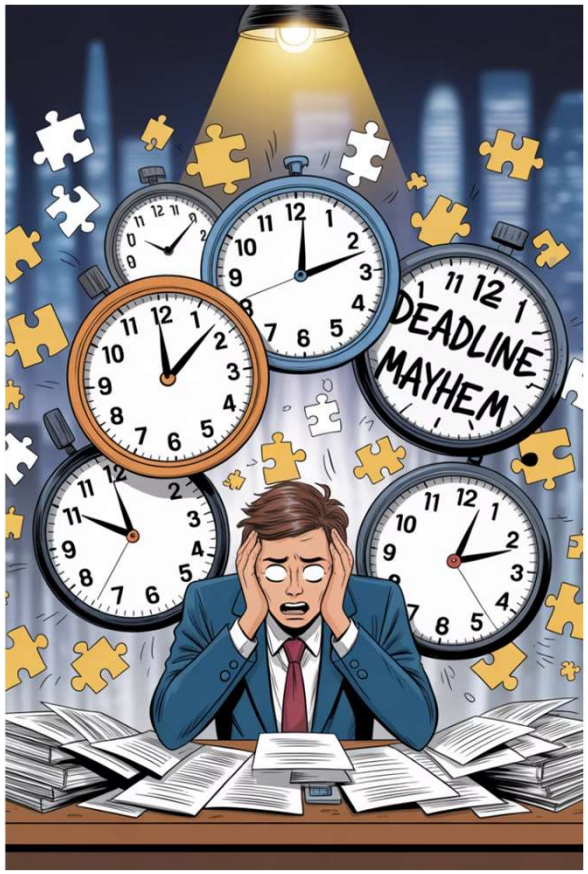
Clinical trials may not capture HTA-relevant endpoints, populations, key subgroups, or sufficient follow-up.

It could lead to post-hoc adjustments (e.g., ML-NMR) fraught with uncertainty, weakening negotiating leverage.

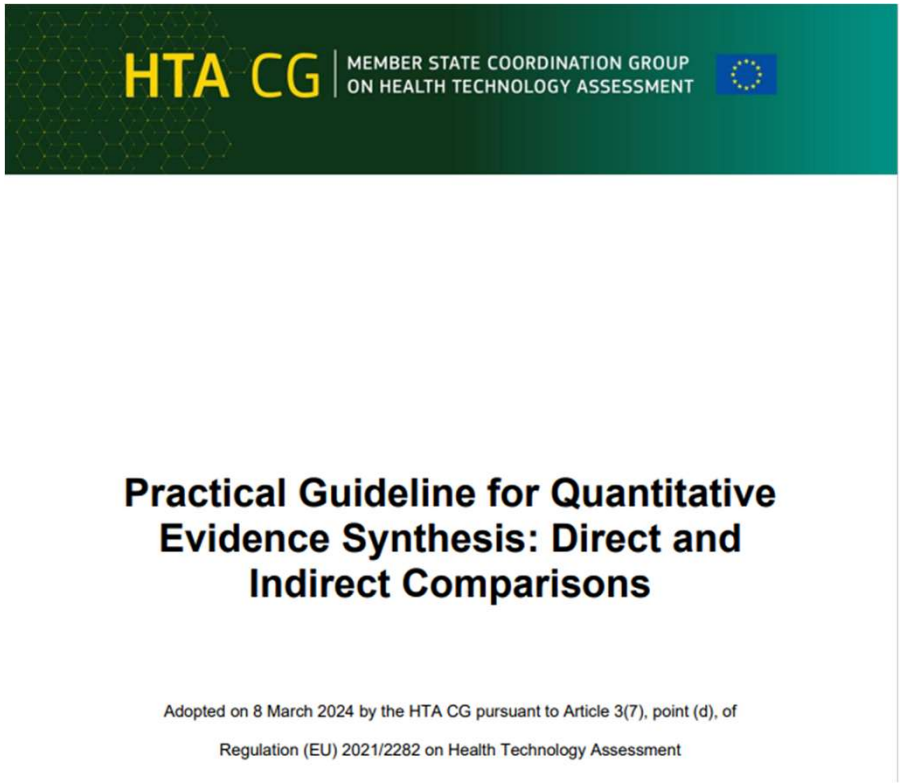
2. Last-Minute Crunch

Rushed ITC work around pivotal readouts is stressful, error-prone and/or lead to suboptimal choices of method.

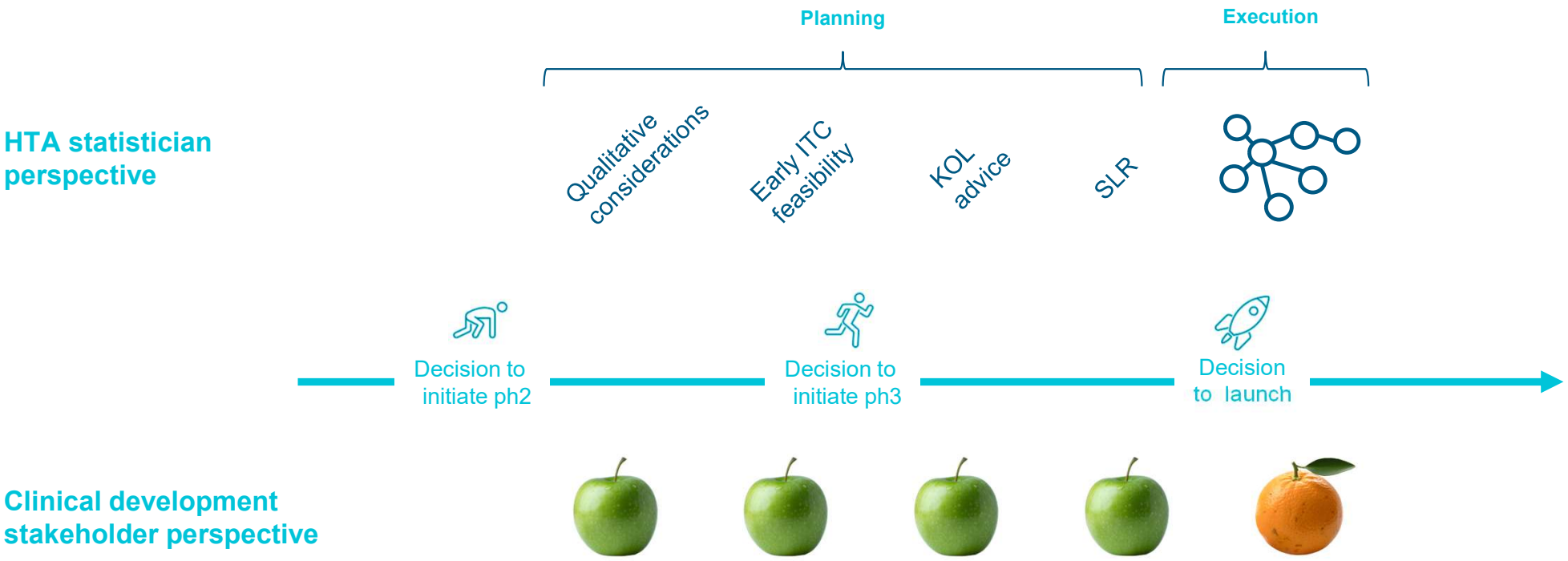
Risks missing critical deadlines



EU HTA prompts us to think about the foundation



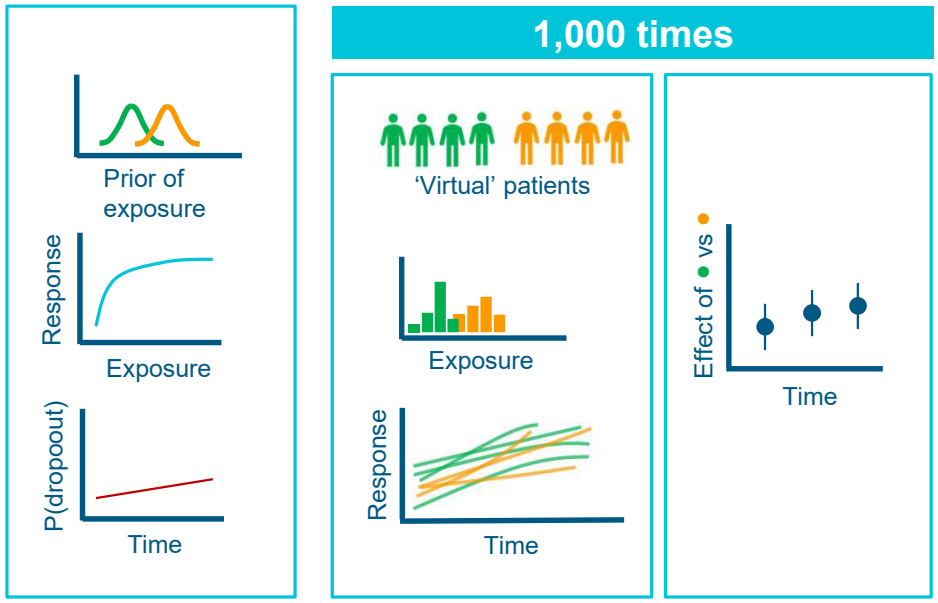
But indirect evidence discussions may not be very accessible to non-technical stakeholders



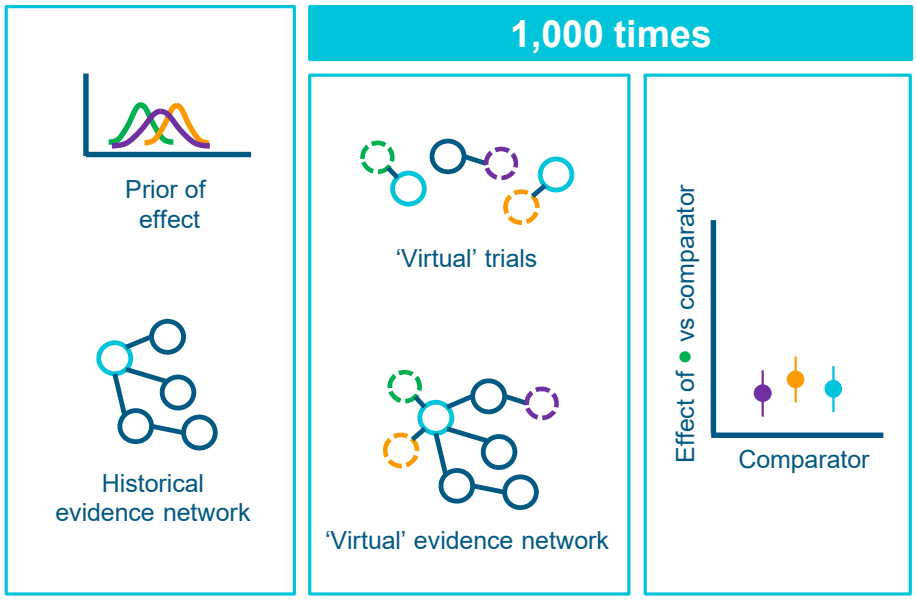
"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

ITC quantitative scenario planning – game changer to drive accessibility of ITC discussions?

CLINICAL TRIAL SIMULATION



'ITC SIMULATION'



"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

Quantitative scenario planning can facilitate stakeholder dialogue & systematic tackling of '***what if...?***' questions

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

Simulated dialogue with Stakeholders

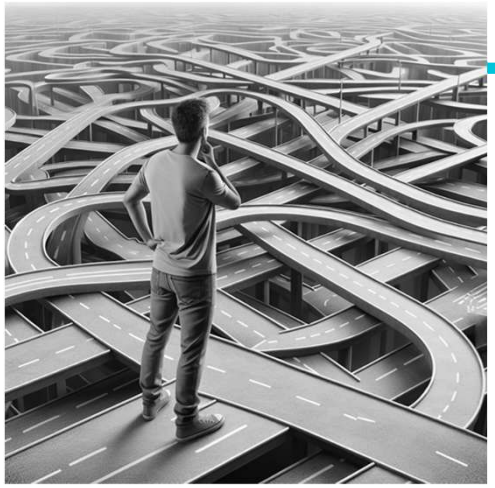
"Do you really need a head-to-head trial – can't you just do an ITC?"

"What if we narrow the trial population wrt. a known effect modifier?"

"Can't you just use [insert favorite innovative statistical method] to fix this?"

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

Why isn’t everyone doing this already?



Resource Constraints

Budgets and headcounts already stretched, leaving little room for additional early-stage work

Uncertain ROI

hard to justify the upfront investment, sunny projections can look cloudy in practice

Complexity Concerns

Juggling extra datasets and cross-functional inputs

Comfort Zone

Established workflows are comfy slippers, and switching to new shoes—even shinier ones—takes time and effort

“We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients.”

A New Opportunity for Statisticians to showcase Leadership, Best Practices, and Tech Savvy?

Proportional Effort

Allocate resources across development stages to focus only on HTA-critical decisions at each step.

Modular Workflows

Break ITC processes into reusable components (data prep, matching, analysis, reporting).

Agile Updates

Implement continuous-integration for data and models—iterate as new information arrives rather than waiting for a full pivot readout.

Simulated Dry-Runs

Run planned ITC methods on simulated datasets well before readouts—so adapting to actual data is straightforward. This is likely needed for mid-stage quantitative scenario planning, one can hit two birds with one stone.

Automation & Templates

Develop scripts, macros, and report templates to reduce manual effort.

Leverage Open-Source

Adopt or contribute to community tools to avoid rebuilding from scratch and stay up to date with best practices.

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

R Consortium HTA working group

Working Group of industry, academia, HTAB participants, inspired by EU HTA JCA challenges and opportunities.

Initial workstreams:

- *Mapping of stakeholders, landscape and opportunities: to strengthen the shared understanding of challenges and potential solutions*
- *Package mapping : develop an inventory of HTA-related R packages related to HTA analytics, facilitating use and integration in daily practice*

https://github.com/RConsortium/HTA-wg

Conclusion: bridging innovation with application in ITCs

- *Less tooling, more schooling*
- Proactive integration of ITC thinking in drug development
- Quantitative scenario analysis as a bridge
- Cross-company collaboration opportunity
- Software engineering as an enabler

EFSPI EUROPEAN FEDERATION OF STATISTICIANS IN THE PHARMACEUTICAL INDUSTRY
Representing Statistical Associations in Europe

Better methods or better planning?

Improving HTA decision-making

Keith Abrams

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."

ML-NMR – Improving HTA decision-making? 1

- What does a methodologist HTA DM look for in a 'new' method?
  - Does it answer an important question/address an important challenge? YES!
  - Does it appear to be complexity for complexity's sake? NO!
  - How does in compare to other methods?
    - Case studies
    - Simulation studies

ML-NMR – Improving HTA decision-making? 2

- What does a HTA DM look for in a ‘new’ method?
  - Does it answer an important question/address an important challenge? YES!
    - We want to compare multiple treatments simultaneously in a target population allowing for population differences in the evidence base.
  - Does it make assumptions which are reasonable? YES!
  - Does it provide results which can be easily understood/used? YES!
  - Does it provide results which other methods fail to do so? YES!

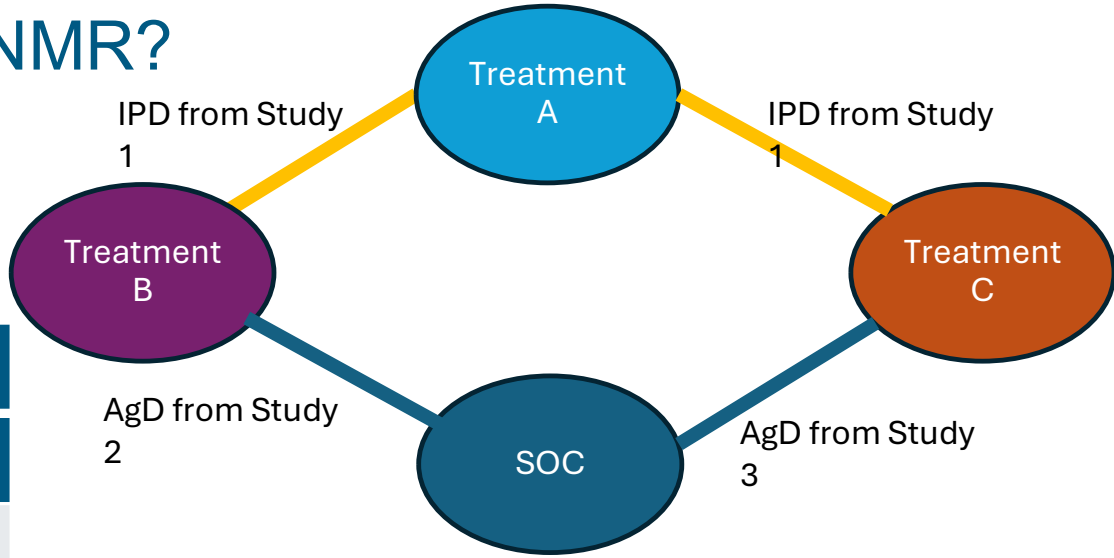
ML-NMR – Improving HTA decision-making? 3

- Consider …
  - "A is as effective as B … C is as effective as D… E is at least as effective as A."
  - "…(is) A as effective as, or inferior to C … (is) E as effective as D or not"
  - "two further questions on indirect comparisons arise, whether C is superior to A or not. and whether E is as effective as C or not."
  - This is actually from a NICE TA on thrombolytic treatment in 2003 before NMA was understood or accepted, BUT it could equally apply to a series of MAICs/STCs in 2025!
  - So MAIC/STC *versus* ML-NMR *cf.* pairwise MA versus NMA

ML-NMR – Improving HTA decision-making? 4

- Are there problems with ML-NMR?
- Consider …

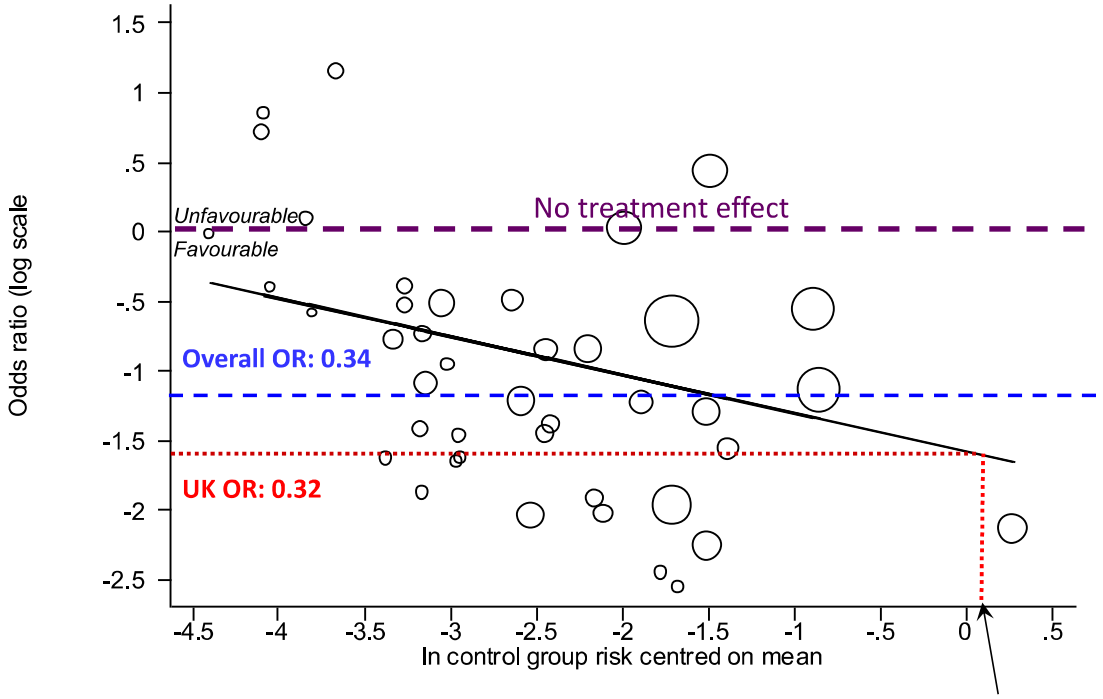
Study/Data	IPD		AgD	
Covariate	A-B	A-C	B-SoC	C-SoC
X_1	✅	✅	✅	✅
X_2	✅	✅	✅	❌
X_3	✅	✅	❌	✅
X_4	✅	✅	✅	✅



Solutions?

ML-NMR – Improving HTA decision-making? 5

- Other approaches to consider …



Baseline-risk/Control/Placebo effect adjusted MA/NMA

- Does not require covariate information
- (Target) Populations are often defined/described by 'risk'
- RTE are often applied to a natural history 'risk' in economic models

Achana FA *et al*. (2013) Extending methods for investigating the relationship between treatment effect and baseline risk from pairwise meta-analysis to network meta-analysis. *Statist Med*,32:752–771. https://doi.org/10.1002/sim.5539

ML-NMR – Improving HTA decision-making? 6

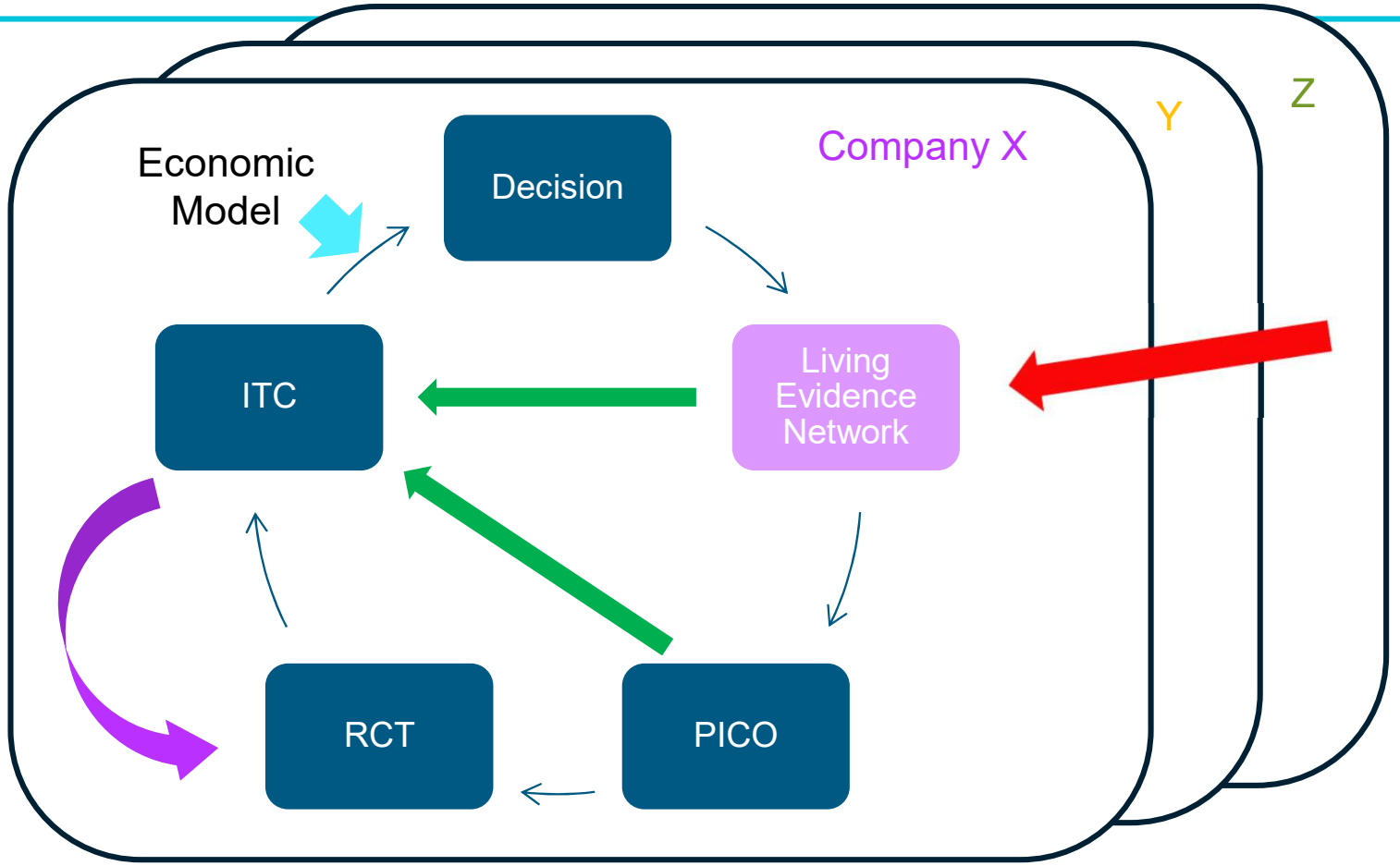
- When an assessors looks at all these ITC results and evidence: What are their thoughts?
  - If you want to convince assessors of what evidence/ITC is required: Is the evidence presented *appropriate* to decision problem and complexity of analyses proportionate?
  - Could we expect differences in NICE/JCA assessment? Possibly (though perhaps less so than previously in EU) – decision space *versus* evidence space, but this should be less crucial if *living evidence network*

ML-NMR – Improving HTA decision-making? 7

- When an assessors looks at all these ITC results and evidence: What are their thoughts?
  - We should know at an early stage where we are going to face issues in the future – this includes in the design of RCTs and what RWD may be required
  - We want to avoid the situation where the ITC should solve problems that could have been addressed or at least identified at an early stage
    - Old RCT adage – always be wary of overly complex analyses!
  - Think of PICOs at an early stage and link these to the ITCs performed (and other studies/data and analyses!)
    - PICO-ITC *cf.* Target Trial Emulation(TTE)-RWD - both are frameworks not solutions!

ML-NMR – Improving HTA decision-making? 8

The *Dynamic* HTA Process



Speakers

Keith Abrams
keith.abrams@warwick.ac.uk

Anders Gorst-Rasmussen
agtr@novonordisk.com

Lytske Bakker
lytskebakker@openhealthgroup.com

Min-Hua Jen
m.jen@lilly.com

Gregory Chen
xiangyi.gregory.chen@msd.com

Nicky Welton
Nicky.Welton@bristol.ac.uk

"We are a community dedicated to leading and promoting the use of statistics within the healthcare industry for the benefit of patients."