19th November 2024

ISPOR EU Educational Symposia

Indirect Treatment Comparison Methodology Matters

Unpacking the Essentials of Robust Analyses

BeiGene connectHEOR University of BRISTOL

Speakers

Shilpi Swami

- Vice President, HTA and Strategy, ConnectHEOR
- Member Engagement Co-Chair at ISPOR Oncology Special Interest Group.

connectHEOR

Rhys Williams

- Vice President, Integrative Evidence Generation and Health Economics, Global Medical Affairs, BeiGene
- Expertise in HEOR, market access, and RWE

BeiGene

Kate Ren

- Director of Statistics, ConnectHEOR
- Senior Research Fellow at University of Sheffield
- Expertise in Statistical Analysis and Evidence Synthesis

connectHEOR

David Phillippo

- Statistician, University of Bristol
- Member of NIHR Bristol Technology Assessment Group.
- Expertise in Bayesian Network meta-analysis and population adjusted methods for indirect treatment comparisons.

University of BRISTOL

Objectives

This educational symposia aims to

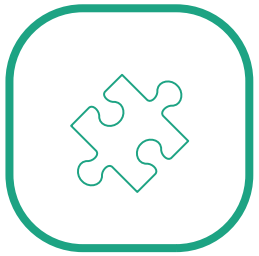
01 Explore diverse ITC methodologies (NMA, MAIC, STC, ML-NMR).

02 Deepen our understanding of each method's strengths, use cases, and limitations.

Why discuss ITC methods?

Importance of ITC

Efficient ITC methods are crucial for fair and applicable treatment comparisons especially as JCA expands across EU.



Challenges in implementation

Many HEOR professional face challenges in determining when and how to use ITC techniques

Need for guidance

Further research and guidance are needed to assist professionals in making informed decision about ITC methods


ITC is
equal to MAIC
CHANGE MY MIND

Poll Time #1!

Which of the following ITC methodologies are you aware of?:

- Network meta-analysis (NMA)
- Matching-adjusted indirect comparison (MAIC)
- Stimulated treatment comparison (STC)
- Multi-level network meta-regression (ML-NMR)

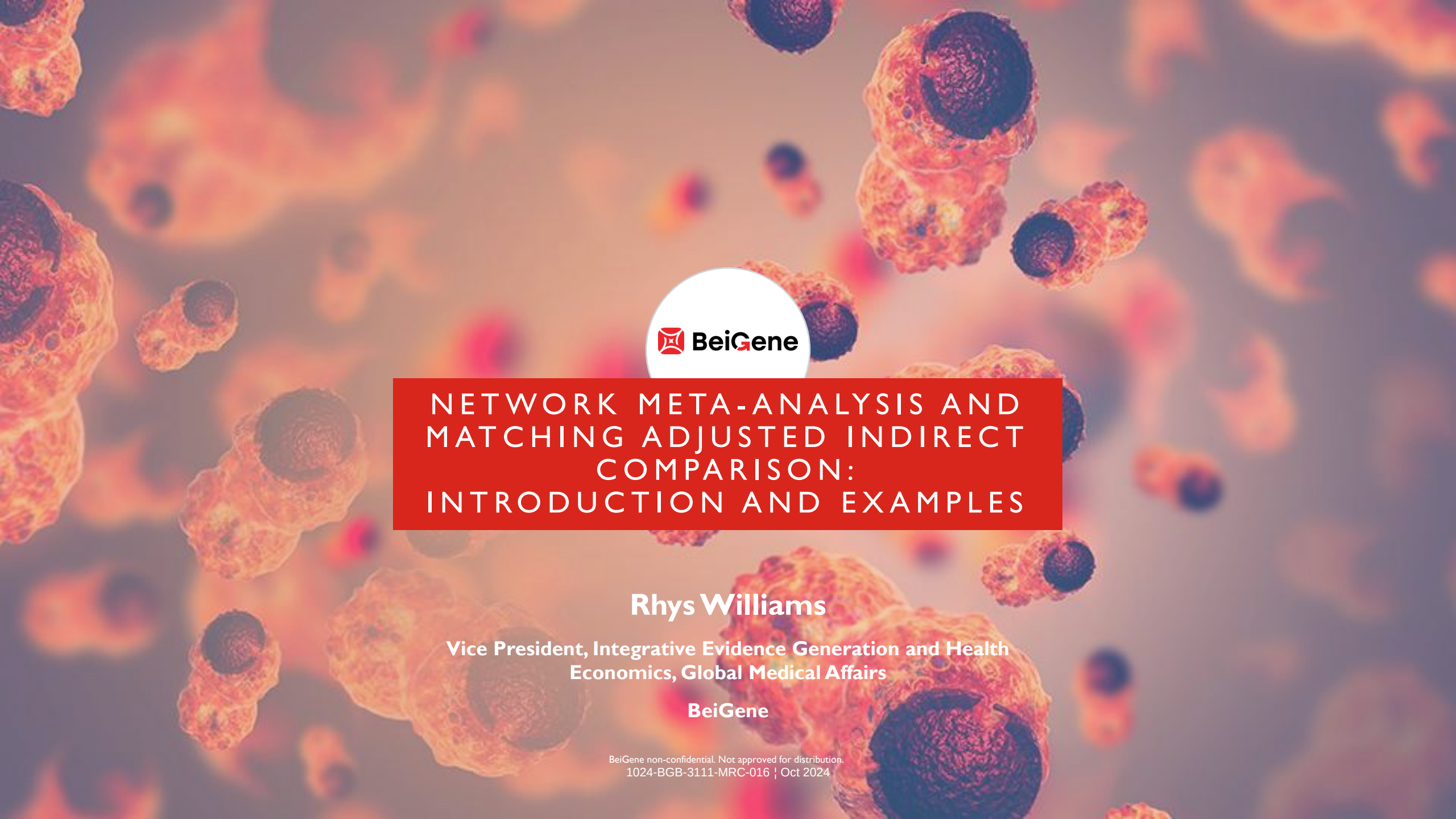
Please select any approaches you've worked with (multiple options allowed).

Poll Time #2!

Which of the following ITC methodologies have you used?:

- Network meta-analysis (NMA)
- Matching-adjusted indirect comparison (MAIC)
- Stimulated treatment comparison (STC)
- Multi-level network meta-regression (ML-NMR)

Please select any approaches you've worked with (multiple options allowed).



NETWORK META-ANALYSIS AND MATCHING ADJUSTED INDIRECT COMPARISON: INTRODUCTION AND EXAMPLES

Rhys Williams

Vice President, Integrative Evidence Generation and Health Economics, Global Medical Affairs

BeiGene

1024-BGB-3111-MRC-016 ¦ Oct 2024

Conflict-of-Interest Statement

- Rhys Williams is an employee of BeiGene and owns stock in BeiGene.

Disclaimers

- The information contained herein is intended for healthcare professionals only and is given for educational purposes only. This document is not intended for professional counseling or advice.
- Matching adjusted indirect comparisons (MAICs) are a methodology to compare data across clinical trials and represent a lower level of evidence than randomized controlled trials.
  - This analysis should be interpreted with caution and should not drive treatment decisions for individual patients
  - Based on the variables selected for matching, outcomes may differ

Research Question: What is the relative efficacy of the BTKis in treatment of R/R CLL?

- Next-generation BTKis have led to changes in the treatment algorithm for patients with high-risk R/R CLL [1]
- Moreover, improved understanding of the CLL genome has facilitated the identification of specific high-risk genetic features of disease, allowing a more personalized approach to treatment [2-6]
- Multiple BTKis are available to treat R/R CLL [7]
- Different methodologies were evaluated[8,9] to estimate relative efficacy of approved and recommended BTKis used to treat R/R CLL

BTKi=Bruton's tyrosine kinase inhibitor, CLL=chronic lymphocytic leukemia, IGHV=immunoglobulin heavy chain variable, NMA=network meta-analysis, R/R=relapsed/refractory.

1. Shadman M. JAMA. 2023;329(11):918-932. 2. Moia R, et al Cancers (Basel). 2020;12(3):642. 3. Hampel PJ, et al Leuk Lymphoma. 2021;62(6):1289-1301. 4. Eichhorst B, et al Ann Oncol. 2021;32(1):23-33. 5. Stephens DM. J Natl Compr Cancer Netw. 2023;21(5.5):563-566. 6. Moia R, et al Expert Rev Hematol. 2020;13(2):109-116. 7. Tam C. *Blood Adv* 2024; 8(9): 2300-2309 8. Shadman M, et al. ASCO 2024 Annual Meeting;abstract 7048. Comparative efficacy of Bruton tyrosine kinase inhibitors in the treatment of relapsed/refractory chronic lymphocytic leukemia: A network meta-analysis (NMA). | Journal of Clinical Oncology (ascopubs.org). 9. Shadman M, et al. Poster Presentation at EHA 2024. https://www.beigenemedical.com/CongressDocuments/Shadman_BGB-3111-305_ASCEND_MAIC_EHA_Poster_2024.pdf

Introduction to Network Meta-Analysis (NMA) and A Case Study Addressing Our Research Question

What is Network-Meta Analysis (NMA)

Any set of studies that links three or more interventions via direct comparisons forms a **network of interventions**. [1]

In a network of interventions there can be multiple ways to make **indirect comparisons** between the interventions. [1]

These are comparisons that have not been made directly within studies, and they can be estimated using mathematical combinations of the direct intervention effect estimates available. [1]

Network meta-analysis combines direct and indirect estimates across a network of interventions in a single analysis.[1]

Example: A network diagram with four competing interventions and two arm and three arm direct comparisons available for some of the trials

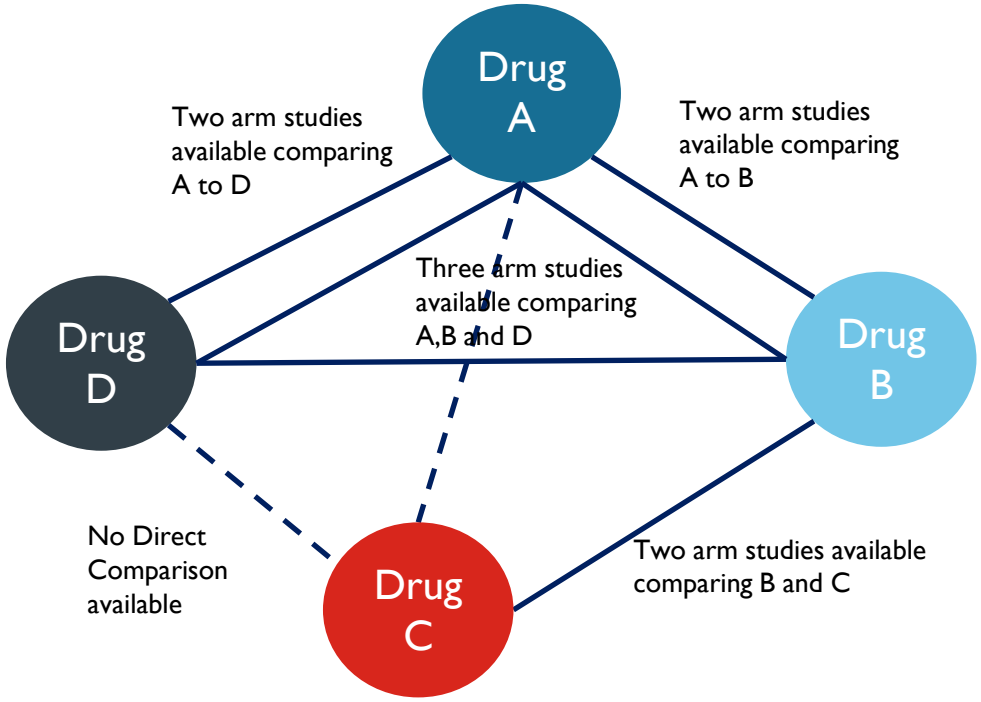


Diagram created by the speaker based on Rouse B, et al. Intern Emerg Med. 2016 Dec 2; 12(1):103-111

1. Rouse B, et al. Intern Emerg Med. 2016 Dec 2; 12(1):103-111

Network Diagram for Our Research Question

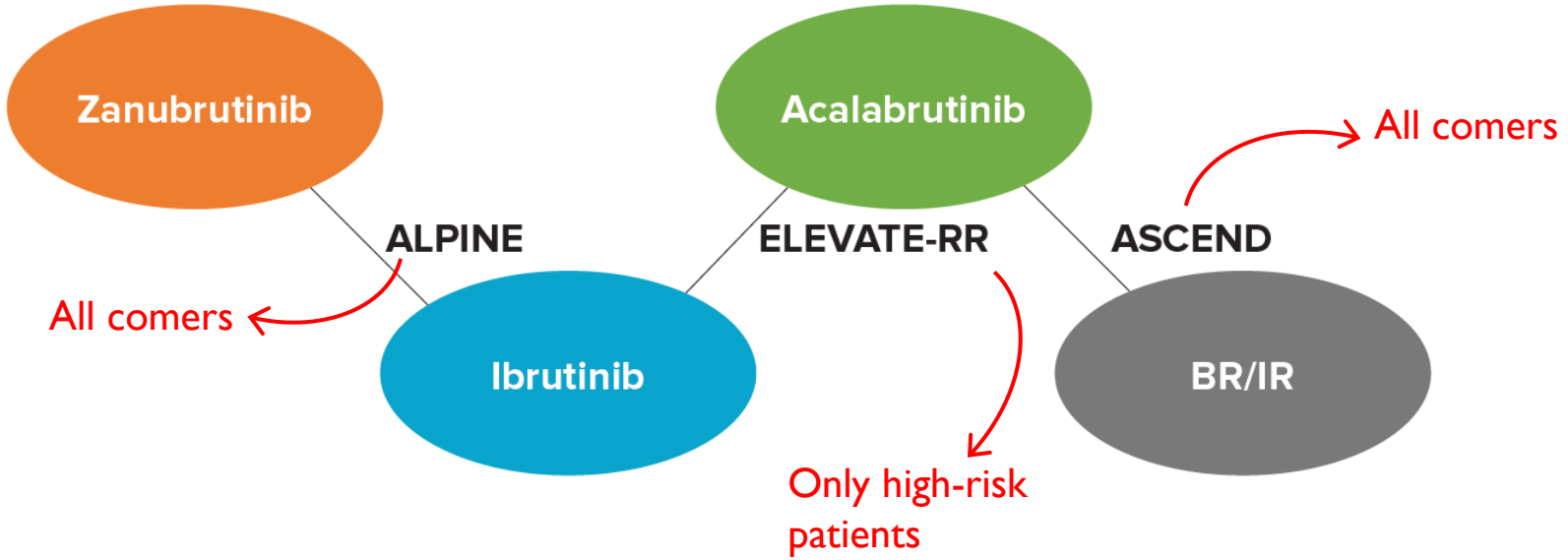


Diagram created by the speaker based on Shadman M et al J Clin Oncol. 2024;42(suppl 16):7048

- When exploring this research question using a Network Meta-Analysis (NMA), the population had to be restricted to only high-risk patients given ELEVATE-RR[1] trial who closes the network includes only high-risk population.
- This reduced the population included from ALPINE[2] and ASCEND[3]
- Given that data from ALPINE were collected during the COVID-19 pandemic, data were analyzed with and without adjustment for COVID-19-related deaths[4]

1. Byrd JC, et al J Clin Oncol. 2021;39(31):3441-3452 2. Brown JR, et al N Engl J Med. 2023;388(4):319-332 3. Ghia P, et al Hemasphere. 2022;6(12):e801 4. Shadman M et al J Clin Oncol. 2024;42(suppl 16):7048

Results

BTKis in R/R CLL: NMA

- In high-risk populations as defined by the individual trials, zanubrutinib was found to be proved significant clinical benefit compared to ibrutinib, acalabrutinib, and BR/IR for progression-free survival, representing risk reductions of 51%, 46%, and 88% respectively, with COVID-19 adjustment

NMA Results Using COVID-19 Adjusted Data from ALPINE Trial – Hazard Ratios and Probability Better for Zanubrutinib vs Comparators

Zanubrutinib vs.	HR [95%CrI]	Probability Better (%)
Acalabrutinib		
PFS	0.54 [0.32, 0.92]	98.6
OS	0.72 [0.35, 1.48]	81.7
Ibrutinib		
PFS	0.49 [0.30, 0.78]	99.9
OS	0.59 [0.31, 1.12]	94.8
BR/IR		
PFS	0.12 [0.05, 0.26]	100
OS	0.64 [0.24, 1.74]	80.7

BR/IR=bendamustine + rituximab or idelalisib + rituximab, CrI=credible intervals, Prob=probability better.
Shadman M, et al. ASCO 2024 Annual Meeting;abstract 7048. Comparative efficacy of Bruton tyrosine kinase inhibitors in the treatment of relapsed/refractory chronic lymphocytic leukemia: A network meta-analysis (NMA). | Journal of Clinical Oncology (ascopubs.org).

Results

BTKis in R/R CLL: NMA

- The table below presents results for zanubrutinib vs acalabrutinib from when data from ALPINE were and were not adjusted for COVID-19 deaths

NMA Results With and Without COVID-19 Adjustment from ALPINE Trial – Hazard Ratios and Probability Better for Zanubrutinib vs Acalabrutinib

Zanubrutinib vs. Acalabrutinib	High-Risk With COVID-19 adjustment	High-Risk With COVID-19 adjustment
HR [95%CrI], Probability Better (%)		
PFS	0.54 [0.32, 0.92], 98.6	0.58 [0.34, 0.98],98.0
OS	0.72 [0.35, 1.48], 81.7	0.84 [0.43, 1.65], 69.1
OR [95%CrI], Probability Better (%)		
ORR	1.91 [0.75, 5.00], 91.7	1.69 [0.61, 4.97], 84.4
CR	2.07 [0.50, 9.67], 84.4	1.84 [0.50, 7.20], 81.6

CR=complete response, CrI=credible intervals,OR=odds ratio; ORR=overall response rate; PFS=progression-free survival
Shadman M et al J Clin Oncol. 2024;42(suppl 16):7048

Limitations

- The definition of high-risk varied between the studies included in this NMA. The ELEVATE-RR trial exclusively enrolled patients with del(17p)/del(11q)[1], while ALPINE[2] and ASCEND[3] did not limit enrollment to this population.
- It is expected that the trials included would have differences in terms of baseline characteristics, however NMAs do not consider any adjustments on the characteristics of the populations.[4]
- The analysis was limited to high-risk R/R CLL patients.[5]

Strengths

- Randomization is preserved with NMAs.[4]
- As no adjustment is made for population characteristics, there is no reduction in sample size as with MAICs.[6]
- This analysis included scenarios with adjustment for the impact of COVID- with results shown to be consistent across different scenarios.[5]

NMA=network meta-analysis; R/R CLL=relapsed/refractory chronic lymphocytic leukemia.

1. Byrd JC, et al J Clin Oncol. 2021;39(31):3441-3452 2. Brown JR, et al N Engl J Med. 2023;388(4):319-332 3. Ghia P, et al Hemasphere. 2022;6(12):e801 4. Watt J, et al. Journal of Investigative Dermatology. 2019;139(1):4-12 5. Shadman M, et al. Poster Presentation at ASCO 2024;abstract 7048 6. Choy E et al. Arthritis Research & Therapy. 2019; 21(32):2019

Introduction to Matching-Adjusted Indirect Comparison (MAIC) and A Case Study Addressing Our Research Question

Matching Adjusted Indirect Comparison

Introduction to MAICs

- MAIC is a statistical method where published estimates of a trial can be combined with Individual Patient Data (IPD) of another trial to obtain indirect estimates[1,2]
- This approach is a form of **propensity score weighting** in which patients in one treatment group (in this case the trial with IPD) are weighted based on their closeness to the other treatment group (in this case the trial with only published aggregate data) [1,2]

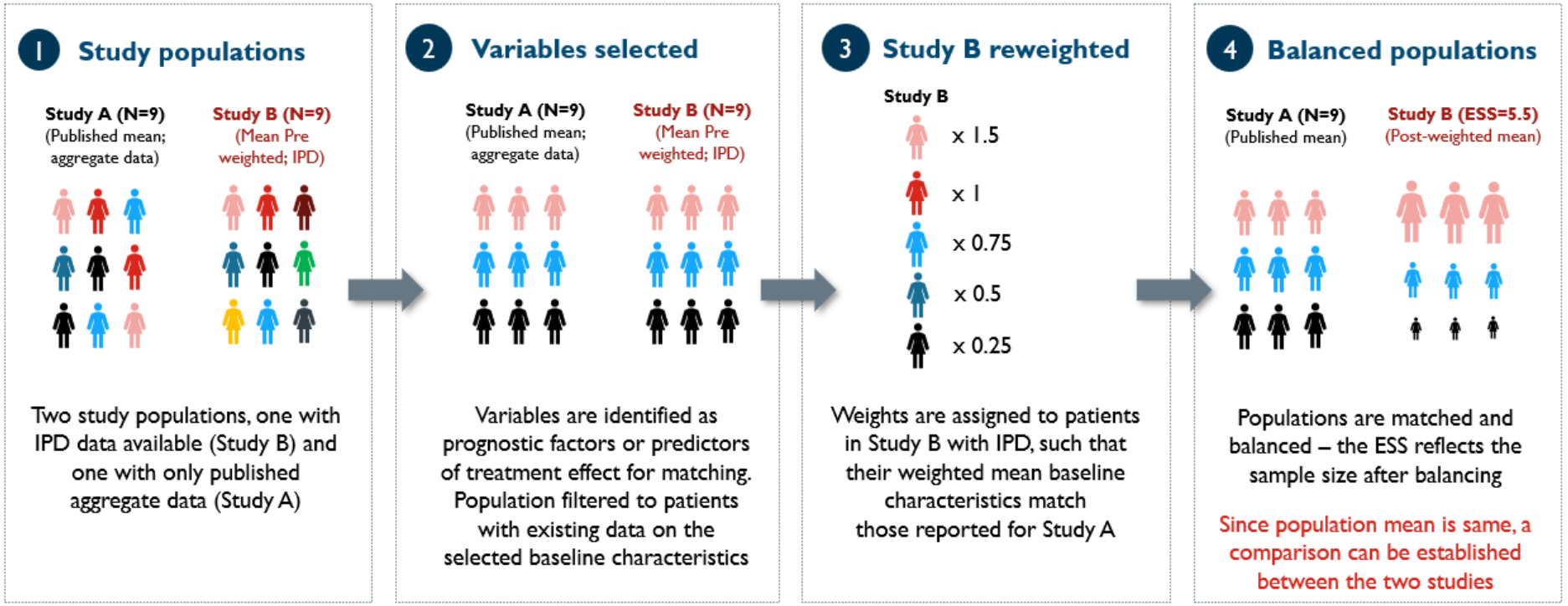


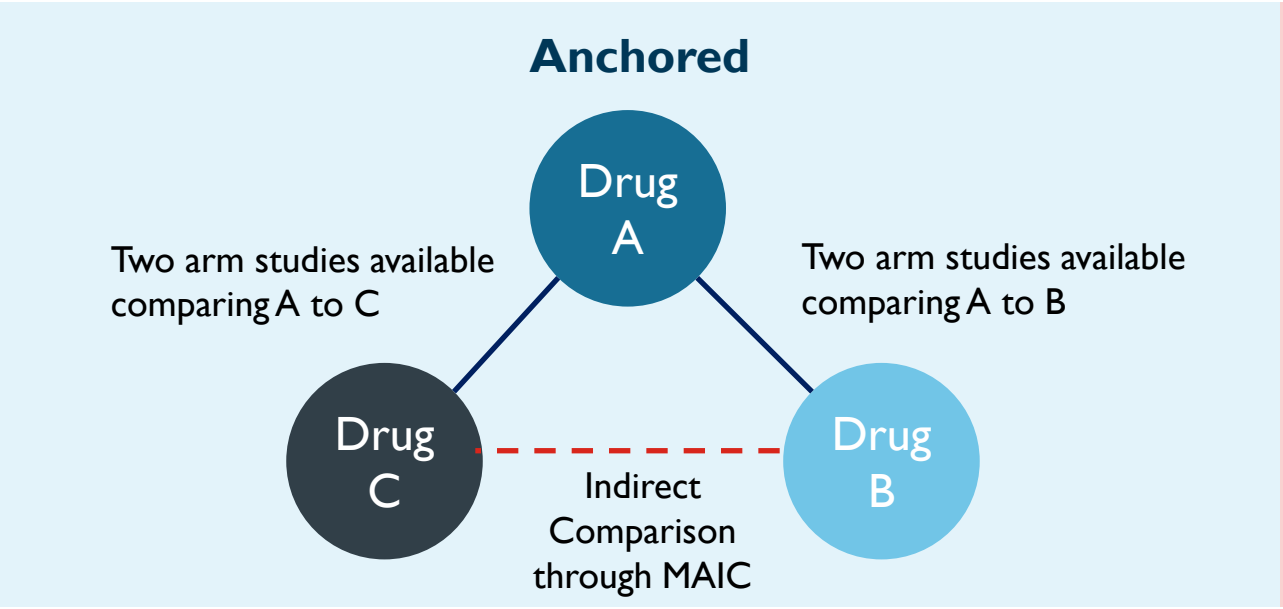
Diagram created by the speaker based on Phillippo[1] and Signorovitch[2]

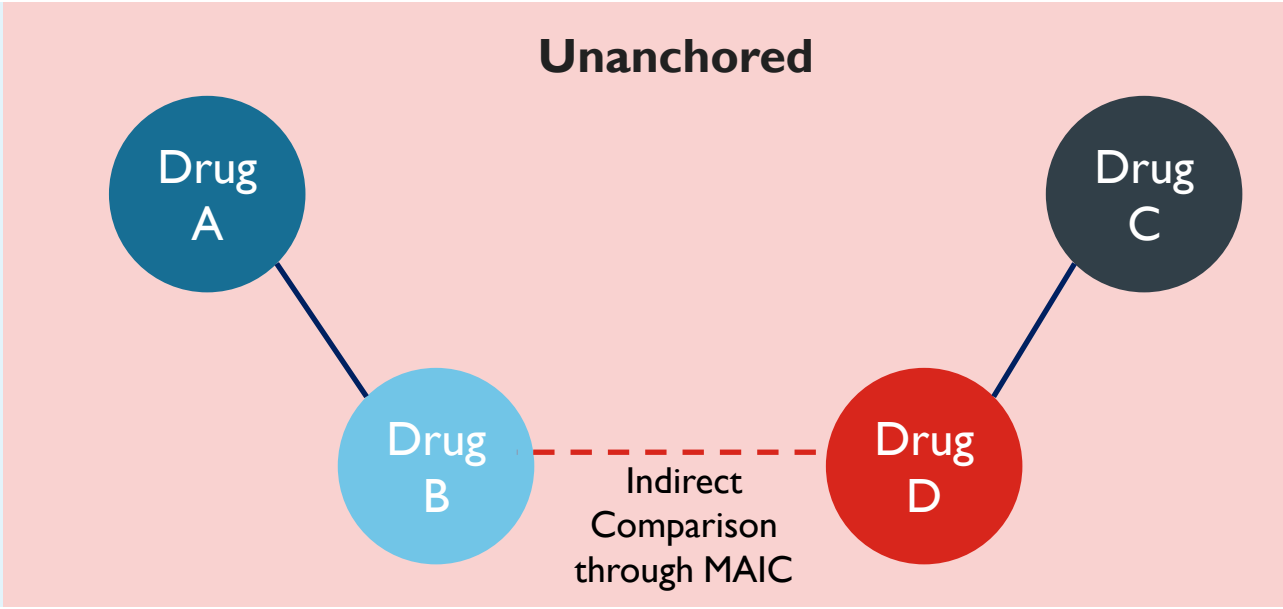
ESS=effective Sample Size, IPD=Individual Patient Data, MAIC=Matching Adjusted Indirect Comparison.

1.Phillippo D, et al. NICE DSU Technical Support Document 18; 2016: https:// research-information.bris.ac.uk/en/publications/nice-dsu-technical-supportdocument-18-methods-for-population-adj; Access date: October 18, 2024 2. Signorovitch JE, et al. Value in Health 2012;15:940-947.

Anchored vs. Unanchored MAIC

Introduction to MAICs



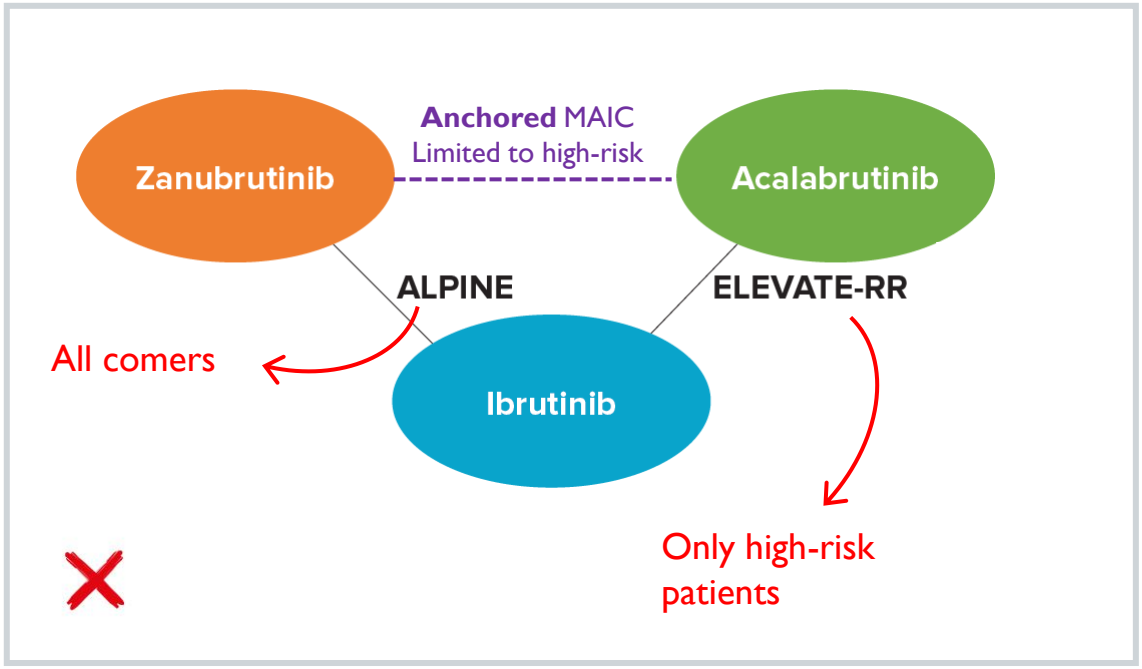


"Anchored" indirect comparisons:
Where the evidence is connected by a common comparator

"Unanchored" indirect comparisons:
Where the evidence is disconnected due to a lack of a common comparator or single-arm studies

MAIC=Matching Adjusted Indirect Comparison.
Phillippo D, et al. NICE DSU Technical Support Document 18; 2016: https:// research-information.bris.ac.uk/en/publications/nice-dsu-technical-supportdocument-18-methods-for-population-adj Access date: October 18, 2024

MAIC options for Our Research Question



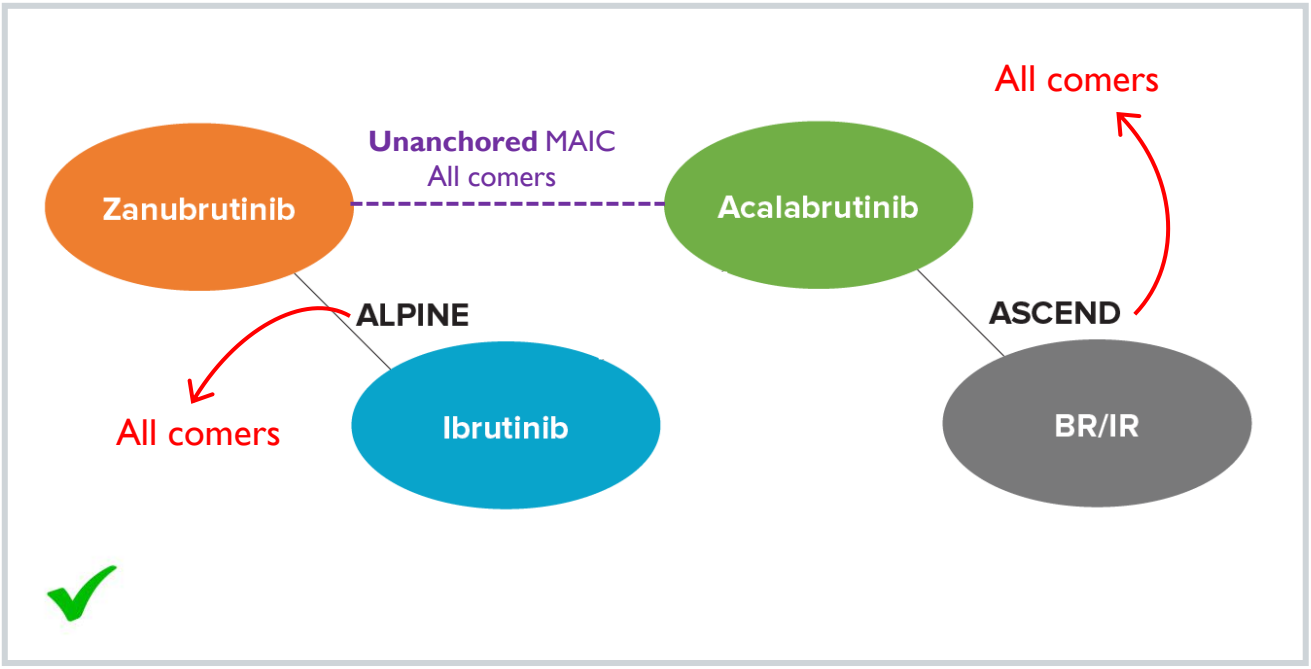


Diagram created by the speaker based on Byrd[1] , Brown [2] and Ghia [3]

- When exploring this research question using a Matching-Adjusted Indirect Comparison (MAIC), two options were possible: anchored and unanchored MAIC
- With an anchored MAIC, the population had to be restricted to only high-risk patients given ELEVATE-RR trial[1] who closes the network includes only high-risk population. This would reduce the starting population. The population would further reduce with matching and adjustment.
- Therefore, an unanchored MAIC comparing zanubrutinib (ALPINE[2]) and acalabrutinib (ASCEND[3]) was preferred.[4]

1. Byrd JC, et al J Clin Oncol. 2021;39(31):3441-3452 2. Brown JR, et al N Engl J Med. 2023;388(4):319-332 3. Ghia P, et al Hemasphere. 2022;6(12):e801 4. Shadman M et al J Clin Oncol. 2024;42(suppl 16):7048

Study Methods

ALPINE vs ASCEND MAIC

ALPINE (N=327)

Individual Patient-Level Data
(DCO: September 2023; median follow-up: 39 months)

ASCEND (N=155)

Published Aggregate Data
(DCO: October 2020; median follow-up: 36 months)

Adjustment for impact of COVID-19 within ALPINE

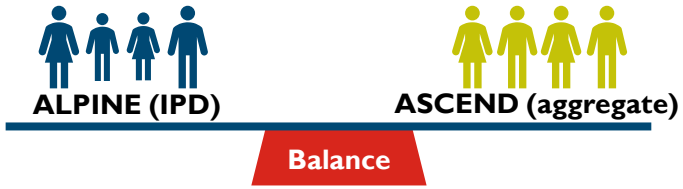
Variable identified as prognostic factors or predictors of treatment effect for matching

Age, gender, ECOG PS, geographic region, mutated IGHV, del(17p), del(11q), *TP53* mutation status, complex karyotype,* bulky disease, cancer type, beta$_2$-microglobulin,* Rai/Binet stage, number and type of prior therapies, absolute lymphocyte and neutrophil counts, and platelet count

Sensitivity analyses of scenarios to consider impact of matching for different sets of variables

Matching, reweighting, and adjusting for variables

- Zanubrutinib unadjusted (ITT) population (ALPINE), **n=327**.
- Zanubrutinib ITT population filtered to patients with existing data on the selected baseline characteristics and excluding patients with SLL, **n=308**.
- After population adjustment, **ESS=184.8** for zanubrutinib (60% of the starting filtered population).



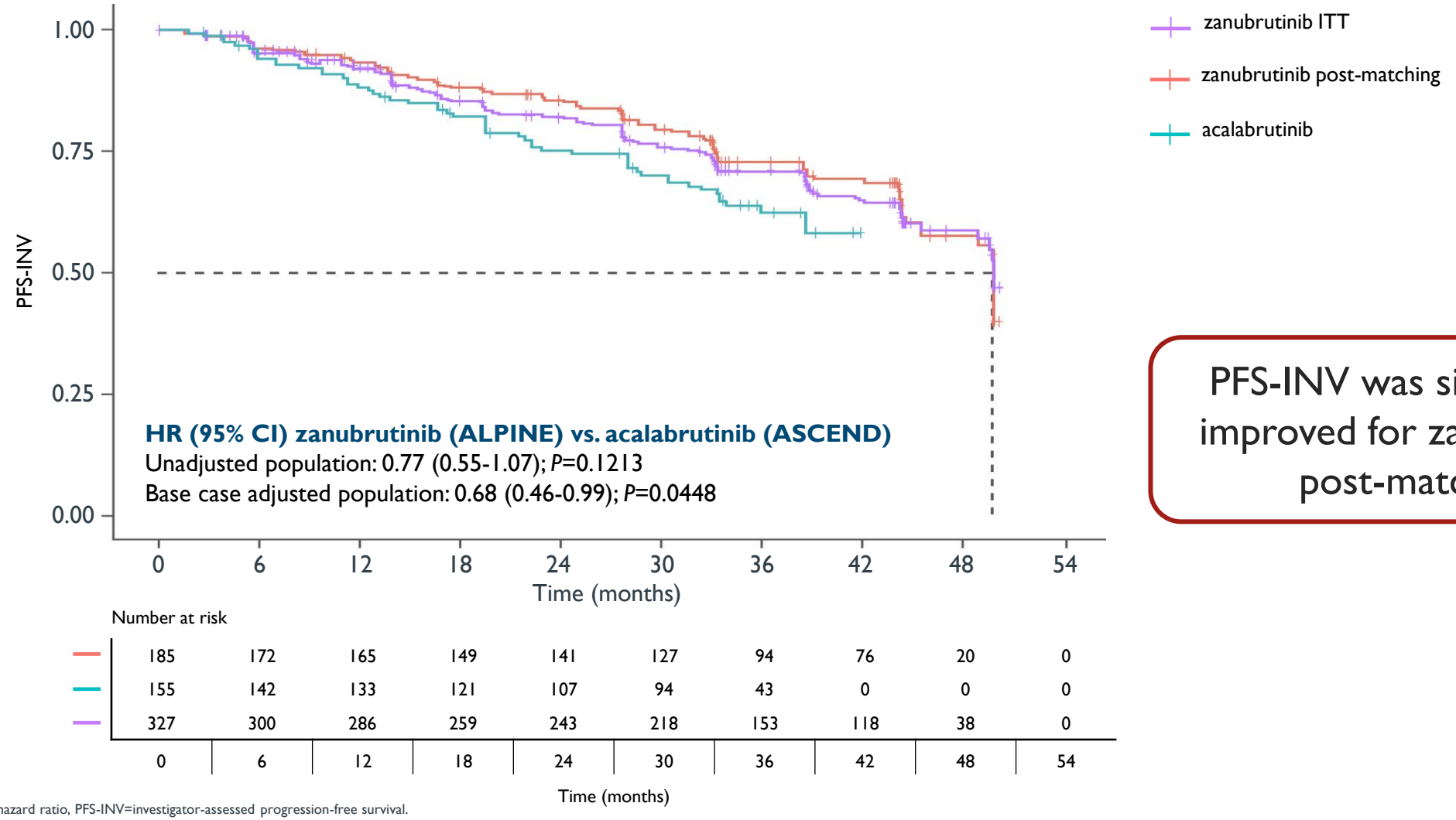
Outcomes

PFS-INV	HRs for PFS-INV and OS: Weighted Cox proportional hazard model
OS	
CR	OR for CR: Weighted logistic regression model

*Covariates not matched in the base case.
CR=complete response, DCO=data cut-off, del(11q)=chromosome 11q deletion, del(17p)=chromosome 17p deletion, ECOG PS=Eastern Cooperative Oncology Group performance status, ESS=effective sample size, HR=hazard ratio, IGHV=immunoglobulin heavy chain variable, IPD=individual patient-level data, OR=odds ratio, OS=overall survival, PFS-INV=investigator-assessed progression-free survival, SLL=small lymphocytic lymphoma.
Shadman M, et al. Poster Presentation at EHA 2024 https://www.beigenemedical.com/CongressDocuments/Shadman_BGB-3111-305_ASCEND_MAIC_EHA_Poster_2024.pdf

PFS-INV for zanubrutinib pre- and post-matching and acalabrutinib

ALPINE vs ASCEND MAIC



PFS-INV was significantly improved for zanubrutinib post-matching

CI=confidence interval, HR=hazard ratio, PFS-INV=investigator-assessed progression-free survival.
Shadman M, et al. Poster Presentation at EHA 2024. https://www.beigenemedical.com/CongressDocuments/Shadman_BGB-3111-305_ASCEND_MAIC_EHA_Poster_2024.pdf

Limitations

- Unanchored MAICs break the randomization and make the strong assumption that cross-trial differences can be entirely explained by variables selected for matching.[1]
- Including all variables in the matching would considerably reduce the ESS, given the differences across the two trial populations and the possibility of missing data on some variables.[1]

Strengths

- Several scenarios with different variables includes and different ESS resulted in consistent conclusions.[2]
- After adjustment, the resulting ESS was rather high given the starting population included the whole ALPINE sample size.[2]

BR/IR=bendamustine + rituximab or idelalisib + rituximab, NMA=network meta-analysis.

1. Signorovitch JE, Value in Health. 2012;15(6):940-947 2. Shadman M, et al. Poster Presentation at EHA 2024. https://www.beigenemedical.com/CongressDocuments/Shadman_BGB-3111-305_ASCEND_MAIC_EHA_Poster_2024.pdf

Differentiation from previous[2] ALPINE vs ASCEND MAIC

ALPINE vs ASCEND MAIC

These results differ from a previously presented MAIC[2] comparing the two trials because limitations of that analysis have been identified and specifically addressed in this analysis – namely:

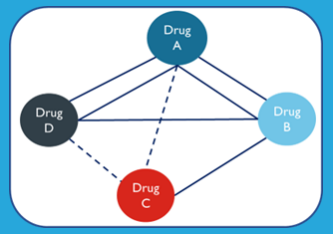
- Selecting comparable published median follow-ups – and the latest analysis of ALPINE
- Accounting for the impact of COVID-19 on ALPINE
- Using clinically-relevant matching criteria
- Ensuring the base case had a large enough sample size – while ensuring adequate sensitivity analyses
- Reporting essential efficacy outcomes
- Leaving safety comparisons for a more robust hypothesis-generating comparison tool – meta-analysis

MAIC=matching-adjusted indirect comparison.

1. Shadman M, et al. Poster Presentation at EHA 2024 https://www.beigenemedical.com/CongressDocuments/Shadman_BGB-3111-305_ASCEND_MAIC_EHA_Poster_2024.pdf; 2. Kittai AS, et al. Am J Hematol. 2023;98:E387–E3

NMA vs. MAIC: Which One to Use?

NMA VS MAIC[1]

	Network Meta Analysis (NMA)	Matching-Adjusted Indirect Comparison (MAIC)
		
1	Compares multiple treatments using published **aggregate data**	Compares published data with **IPD** (Individual Patient Data)
2	Requires a common comparator (connected evidence network, controlled trials)	May not require a common comparator (unanchored comparisons, single arm trials)
3	Assumes trials are comparable in terms of design and population (heterogeneity cannot be handled)	Some of the heterogeneity can be handled between trials by matching the patient population
5	Traditional and established methodology	Evolving method
6	There is no loss of information in methodology as all the available information is used	In an attempt to handle the heterogeneity, there is a risk of reduction of patient sample size

MAIC=matching-adjusted indirect comparison; NMA=network meta analysis
1. Extracted from Choy E et al Arthritis Res Ther. 2019;21:32

Feel free to contact us for any queries

rhys.williams@beigene.com

connectHEOR

Unanchored Simulated Treatment Comparison (STC) for Time-to-Event Outcomes

Speaker:

Kate Ren, PhD

Director, Statistics | Senior Research Fellow

ConnectHEOR, London, UK | University of Sheffield, UK

Disclaimer

- The development of the unanchored STC methodology was funded by the National Institute for Health and Care Research (NIHR) Advanced Fellowship Program.

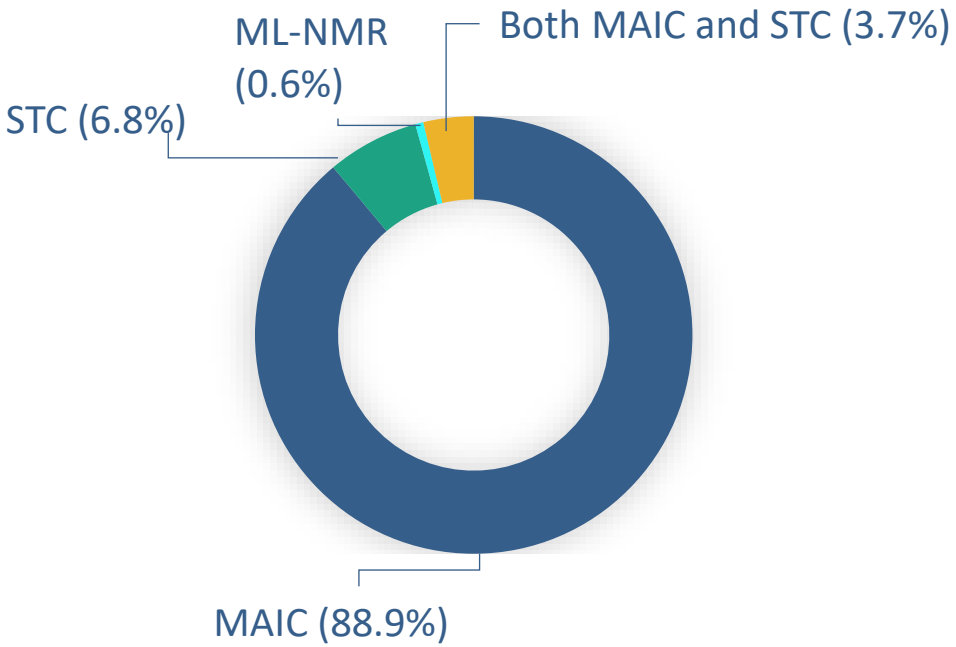
- ConnectHEOR supported the conduct of the case study presented in this talk.

- The views and content of the presentations are the responsibility of the speaker.

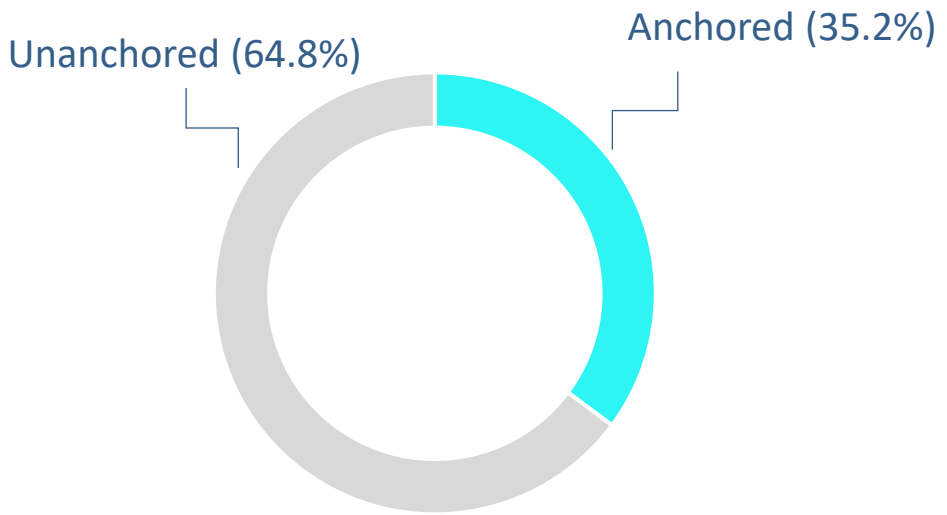
Background

- Population-adjusted indirect comparisons (PAICs) are increasingly used to adjust for the population differences between trials in HTA
- A methodological systematic review[1] of studies implementing PAICs shows that

Population Adjusted Methods (%)



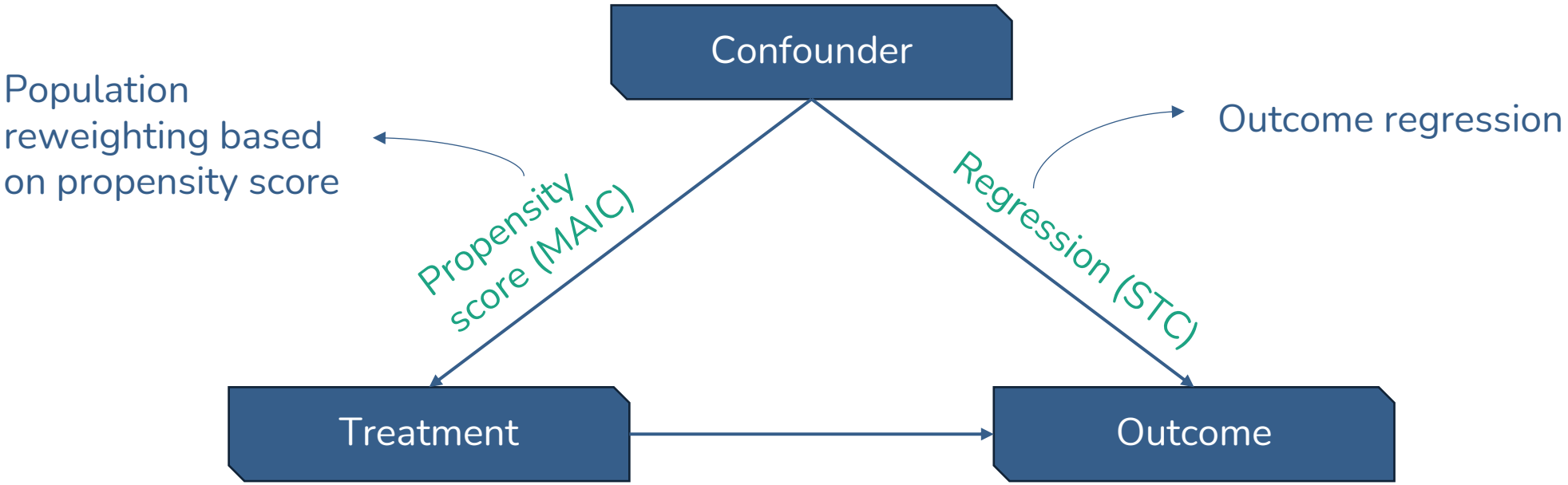
Types of Comparison (%)



[1] Truong et al. (2023) Population adjusted-indirect comparisons in health technology assessment: a methodological systematic review. Research Synthesis Methods

Abbreviations: PAICs: population-adjusted indirect comparisons; HTA: health technology assessment; MAIC: matching-adjusted indirect comparison; STC: simulated treatment comparison; ML-NMR: multilevel meta-regression

Unanchored population adjustment



For both methods, all effect modifiers and prognostic factors should be adjusted for.

Abbreviations: MAIC: matching-adjusted indirect comparison; STC: simulated treatment comparison; ML-NMR: multilevel meta-regression

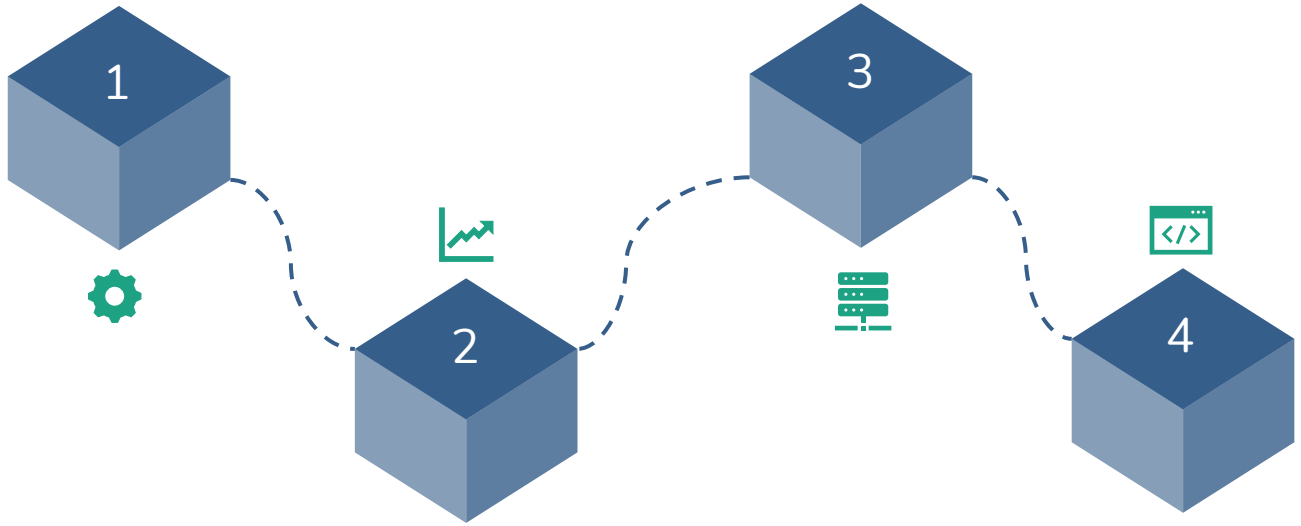
MAIC

MAIC is based on propensity score weighting

Unanchored MAIC steps include:

1. Derivation of a logistic propensity score model based on the IPD from the manufacturer's trial, including all effect modifiers and prognostic factors. Weights are often estimated using the method of moments.

3. Derivation of the indirect comparison in the AgD study population.



2. Prediction of outcomes for the AgD study population by reweighting the outcomes of the IPD study individuals.

4. Calculation of standard error using robust sandwich estimator.

Ref: Signorovitch JE, Sikirica V, Erder MH, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. Value Health. 2012; 15(6): 940-947. doi:10.1016/j.jval.2012.05.004

connectHEOR

Abbreviations: AgD: Aggregated data; IPD: Individual patient-level data; MAIC: matching-adjusted indirect comparison

Matching-adjusted indirect comparison

Strength

01 MAIC is the most frequently used population adjustment method.

02 MAIC's reweighting procedure to balance population differences is intuitive and given the reweighted data, standard statistical analysis methods can be used to obtain the indirect treatment effect.

03 In MAIC, weights were only calculated the weights once. The estimated weights can then be used in the indirect treatment comparison for all outcomes measures.

04 MAIC always estimates a marginal or population-average treatment effect (i.e. the average effect at the population level), which is the required estimate in HTA.

Limitations

01 MAIC does not work well in the case with limited overlap, because of the large reduction in the effective sample size after weighting.

02 In an extreme case of a lack of covariate overlap (large number of covariates and small sample size) MAIC may fail to produce feasible weights.

03 MAIC can only adjust for observed covariates. The unmeasured confounding may still bias the results.

04 MAIC can only produce estimates of quantities of interest in the AgD study population.

05 MAIC only works for pairwise comparison.

Abbreviations: AgD: Aggregated data; HTA: health technology assessment; MAIC: matching-adjusted indirect comparison

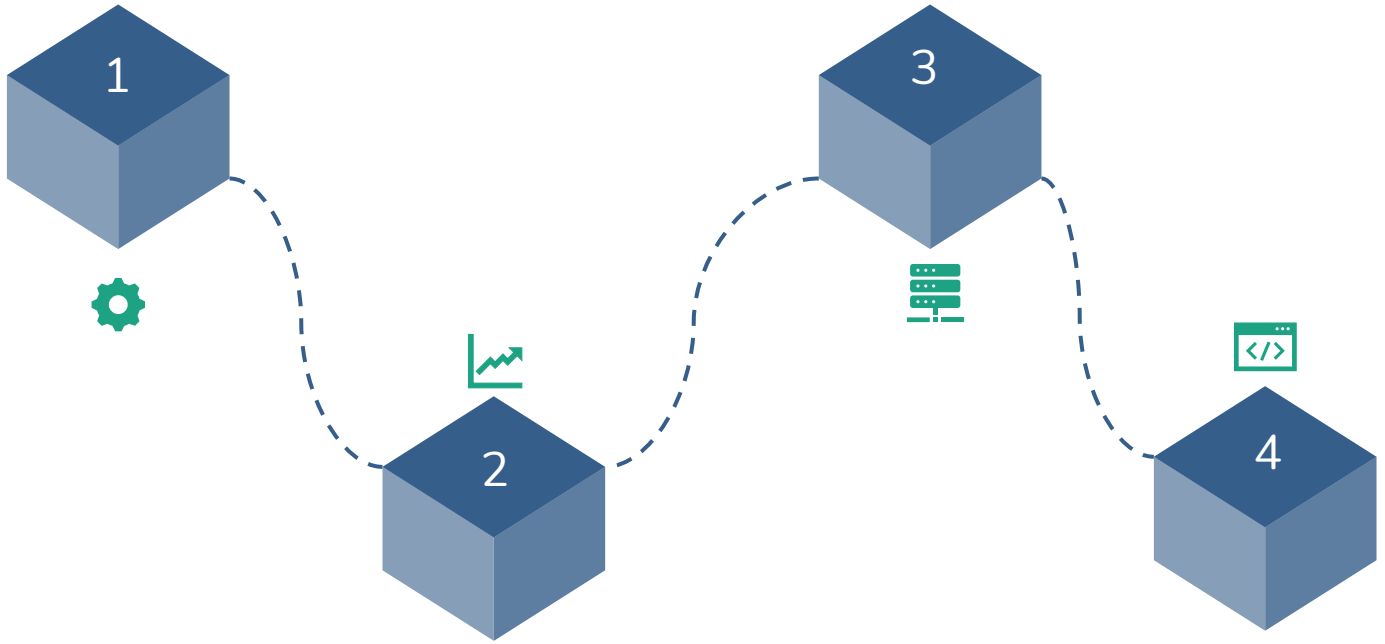
Unanchored STC steps include:

STC

STC is an outcome regression-based modelling approach, which relies on regression models with covariates included

1. Development of a regression model based on the IPD from the IPD study, including all effect modifiers and prognostic factors.

3. Derivation of the indirect comparison in AgD study population, using the prediction from Step 2 and reported aggregate data for AgD study.



2. Prediction of the outcome for AgD study population based on the regression model from Step 1.

4. Calculation of standard error using bootstrapping.

Ref: Caro JJ, Ishak KJ. No head-to-head trial? Simulate the missing arms. Pharmacoeconomics. 2010; 28(10): 957-967. doi:10.2165/11537420-000000000-00000

connectHEOR

Abbreviations: AgD: Aggregated data; IPD: Individual patient -evel data; STC: simulated treatment comparison

Simulated treatment comparison

Strength

01 STC approach would be more efficient than MAIC approach because regression-based approaches give more precise estimates, under correct model specification.

02 This advantage becomes more notable where overlap is poor between the studies. MAIC may fail to produce feasible weights. STC would still be feasible due to its ability to extrapolate beyond the observed covariate space in the IPD study.

03 Model assumption could be checked explicitly.

04 Unanchored STC always estimates a marginal or population-average treatment effect (i.e. the average effect at the population level), which is the required estimate in HTA.

Limitations

01 STC is less intuitive to understand compared to MAIC.

02 STC requires a correct specification of the regression model. Different outcomes require different regression model.

03 STC can only adjust for observed covariates. The unmeasured confounding may still bias the results.

04 STC can only produce estimates of quantities of interest in the AgD study population.

05 STC only works for pairwise comparison.

06 Anchored STC can produce either conditional or marginal treatment effect. Careful consideration is required to make sure the correct estimand is obtained.

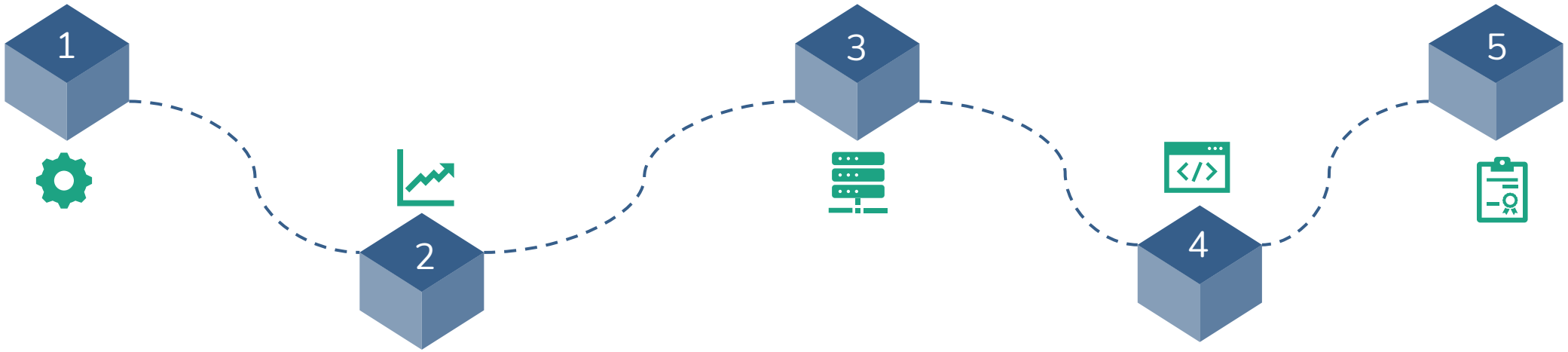
Abbreviations: AgD: Aggregated data; HTA: health technology assessment; STC: simulated treatment comparison

Unanchored STC approach for TTE - procedures

1. Fit outcome regression model
Fit a cox regression model to the IPD study data with prognostic factors and effect modifiers as covariates.

3. Predict survival probabilities for AgD population
The survival probabilities when the AgD population receiving new treatment can be predicted with based on the outcome regression in Step 1 and the simulated covariates in Step 2.

5. Obtain the relative treatment effect
Given the reconstructed IPD from Step 4, standard survival extrapolation methods can be applied as usual. Use bootstrap method to estimate the standard error.



2. Simulate covariates for AgD study
Sample individual-level covariates for the AgD population using Gaussian copula given the published summary statistics of the patient characteristics.

4. Reconstruct IPD for both arms in the aggregate population
Using Guyot's method to reconstruct

- the IPD of the aggregate population receiving new treatment from the predicted survival probabilities from Step 3.
- the IPD of the aggregate population receiving control treatment from the published Kaplan-Meier curves.

Abbreviations: AgD: Aggregated data; IPD: Individual patient level data; STC: simulated treatment comparison; TTE: time-to-event

Case study - Mantle cell lymphoma

Mantle cell lymphoma (MCL) is a distinct subtype of B-cell non-Hodgkin lymphoma.

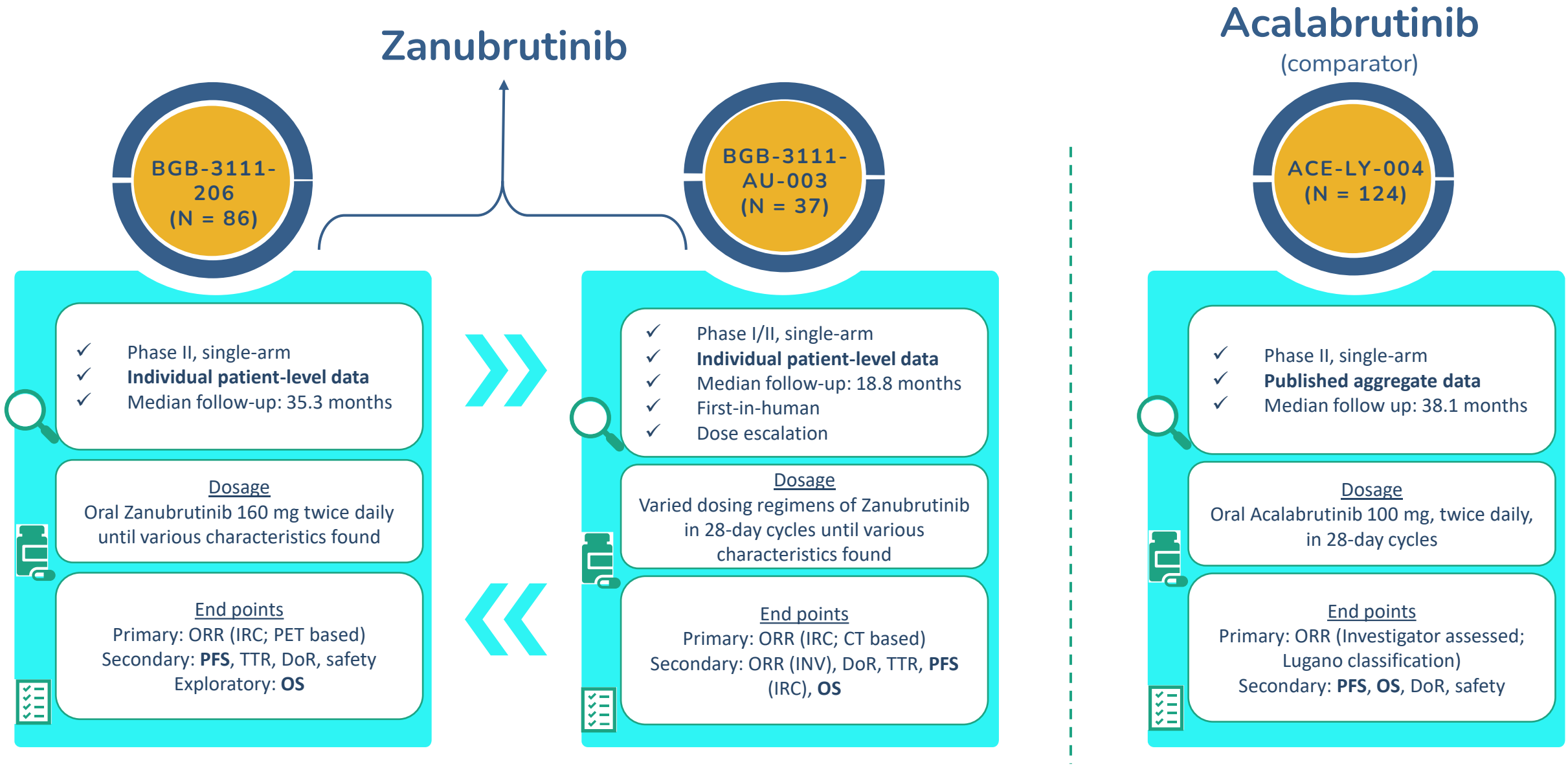
- MCL: ~5% of all non-Hodgkin lymphomas
- **Median age at diagnosis** ranges from 60 to 70 years
- **Median overall survival** is reported to be 3–5 years

Treatments for relapsed/refractory MCL

- Zanubrutinib
- Acalabrutinib

Lack of **comparative evidence** between zanubrutinib and acalabrutinib

Zanubrutinib and acalabrutinib trials from TLR



connectHEOR

Abbreviations: CT: computed tomography; DoR: duration of response; INV: investigator-assessed; IRC: independent review committee; ORR: overall response rate; OS: overall survival; PET: positron emission tomography; PFS: progression-free survival; TTR: time to response; TLR: targeted literature review

Baseline characteristics

	Zanubrutinib	Acalabrutinib
Baseline characteristics (proportion of patients)	**BGB-3111-206 + AU003 (Pooled N=123)**	**ACE-LY-004 (N=124)**
Age ≥65 years	39.8%	64.5%
Race: White	24.4%	74.2%
Sex: Male	74.8%	79.8%
ECOG PS 1–2 (vs. 0)	35.8%	42.7%
sMIPI intermediate risk (vs. low)	37.4%	43.9%
sMIPI high risk (vs. low)	15.4%	17.1%
Bulky disease (LD ≥5 cm)	38.8%	37.1%
Ann Arbor stage III–IV	90.2%	75.0%
Extranodal disease	57.7%	72.6%
Lactate dehydrogenase, high	38.2%	26.6%
Prior lines of treatment >2	32.5%	22.6%
Bone marrow involvement	49.6%	50.8%
Prior autologous SCT	8.9%	17.7%

Noticeable differences observed in several baseline characteristics between the acalabrutinib and zanubrutinib study populations

Note

Indirect treatment comparison without any adjustment would lead to biased results.

connectHEOR

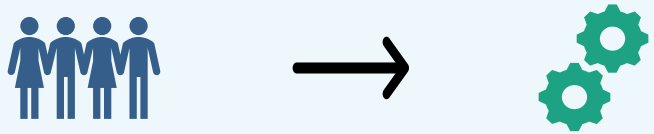
Abbreviations: ECOG: Eastern Cooperative Oncology Group; sMIPI: Simplified Mantle Cell Lymphoma International Prognostic Index; SCT: stem cell transplant

Unanchored MAIC

- Limited overlap
- Substantial reduction in effective sample size after weighting
- Only adjusting for a subset of covariates
- ? Robustness of the analysis

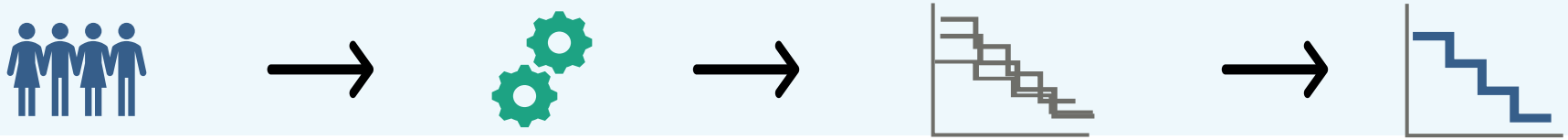
Unanchored STC: PFS and OS

Step 1: FIT REGRESSION MODEL



Zanubrutinib data | Cox regression model

Step 2: PREDICT OUTCOME



Zanubrutinib data | Cox regression model | Predict the survival probability for acalabrutinib trial population | Marginalisation to obtain the overall effect for acalabrutinib trial population treated with zanubrutinib

Step 3: COMPARE RESULTS

Zanubrutinib data | Cox regression model | Reconstruct the IPD based on the predicted KM from Step 2 and reported KM from acalabrutinib trial

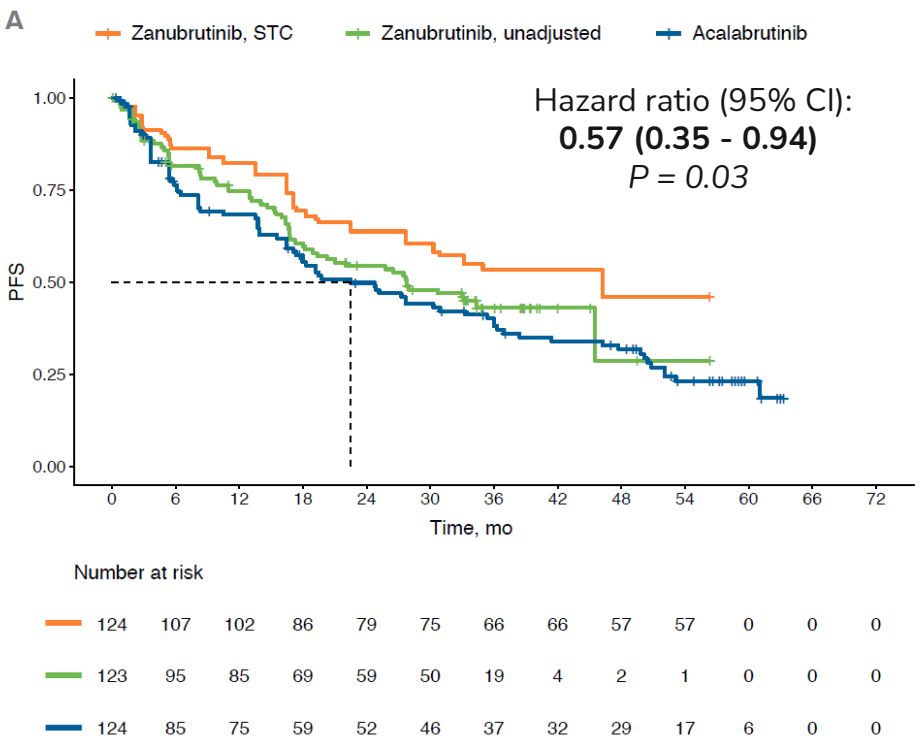
Acalabrutinib trial population treated with **zanubrutinib**

Acalabrutinib trial population treated with **acalabrutinib**

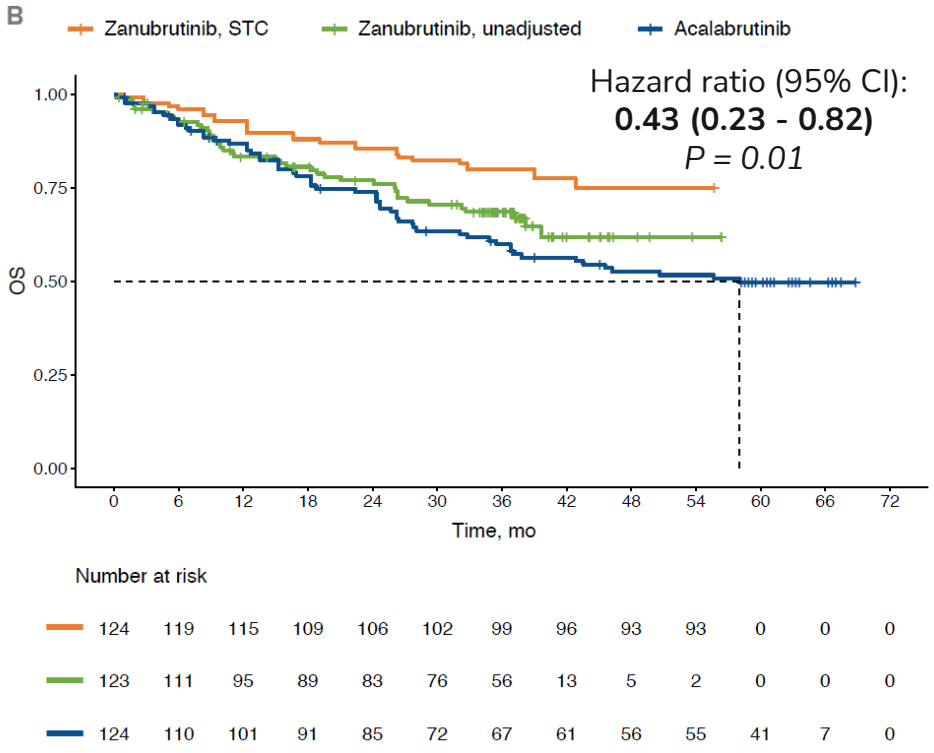
Abbreviations: STC: simulated treatment comparison; PFS: progression-free survival; OS: overall survival

Results

Kaplan-Meier curve for progression-free survival



Kaplan-Meier curve for overall survival



	PFS	OS
Sensitivity analysis 1 (base case without race)	**0.62 (0.39-0.98)** P=0.04	**0.42 (0.25-0.70)** P<0.01
Sensitivity analysis 2 (base case without age)	**0.58 (0.35-0.97)** P=0.04	**0.48 (0.25-0.94)** P=0.03

The sensitivity analyses provided results consistent with base case analysis, demonstrating the robustness of the STC analysis.

Abbreviations: STC: simulated treatment comparison; PFS: progression-free survival; OS: overall survival; CI: confidence interval

Summary

1. Unanchored MAIC ---> robust ?
   Substantial reduction in effective sample size

2. Unanchored STC allows to adjust for all observed covariates.

3. Unanchored STC shows that treatment with zanubrutinib was associated with greater PFS and OS vs. acalabrutinib in R/R MCL.

Abbreviations: STC: simulated treatment comparison; MAIC: matching-adjusted indirect comparison; PFS: progression-free survival; OS: overall survival; R/R: relapse or refractory; MCL: mantle cell lymphoma;

Concluding remarks

Issue

MAIC does not work well in the case with limited overlap.

What have we done?

Demonstrate a novel way to implement unanchored STC for TTE.

Take-home message

- **STC** should be considered for population-adjusted indirect comparisons.
- Care needs to be taken in the implementation to ensure
  - the derivation of unbiased estimate for the marginal treatment effect
  - appropriately quantified uncertainty associated with it.

Abbreviations: STC: simulated treatment comparison; MAIC: matching-adjusted indirect comparison; PFS: progression-free survival; OS: overall survival; R/R: relapse or refractory; MCL: mantle cell lymphoma;

References

- Ren S, Ren S, Welton NJ, Strong M. Advancing unanchored simulated treatment comparisons: A novel implementation and simulation study. Res Syn Meth. 2024; 15(4): 657-670. doi:10.1002/jrsm.1718
- Truong B, Tran LT, Le TA, Pham TT, Vo TT. Population adjusted-indirect comparisons in health technology assessment: A methodological systematic review. Res Synth Methods. 2023 Sep;14(5):660-670. doi: 10.1002/jrsm.1653. Epub 2023 Jul 3. PMID: 37400080.
- Signorovitch JE, Sikirica V, Erder MH, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. Value Health. 2012; 15(6): 940-947. doi:10.1016/j.jval.2012.05.004
- Caro JJ, Ishak KJ. No head-to-head trial? Simulate the missing arms. Pharmacoeconomics. 2010; 28(10): 957-967. doi:10.2165/11537420-000000000-00000

Thank you!

Kate.ren@connectheor.com

W: www.connectHEOR.com

E: info@connectheor.com

Multilevel network meta-regression for time-to-event outcomes

David Phillippo

Bristol Medical School (Population Health Sciences)

University of Bristol, UK

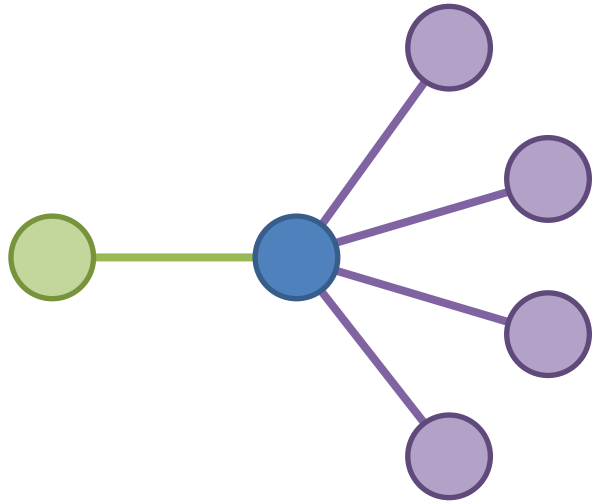
University of BRISTOL

Declarations

- Received speaker fees and expenses from BeiGene
- I will show a BeiGene case study, in which I had no involvement

Multiple comparators – problematic for MAIC and STC

- Larger networks are already commonplace in HTA
  - 2019 review of NICE TAs with population-adjustment found 56% involved larger networks
- Likely to increase with JCA
  - Required to consider more comparators
- MAIC and STC cannot handle larger networks
  - Multiple analyses are incoherent, re-use the data
  - Each analysis valid for a different target population



Further limitations of MAIC and STC

- Comparisons are stuck in the aggregate study population
  - May not be relevant target population for decision making
- STC can incur aggregation bias with non-linear models, non-collapsibility bias
  - Must use appropriate simulation/g-computation methods
- Key assumptions cannot be assessed
  - Conditional constancy of relative effects – unobserved effect modifiers (EMs)

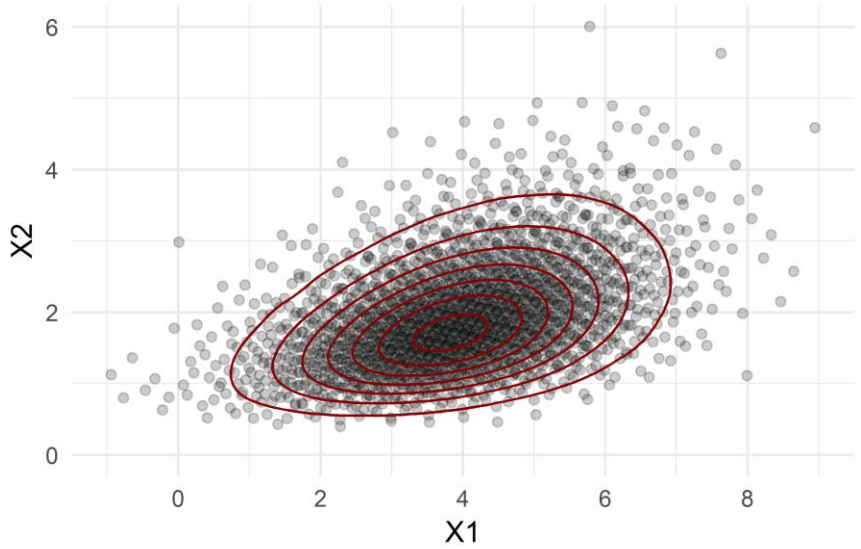
Multilevel Network Meta-Regression (ML-NMR)

Phillippo et al. (2020)

- Applicable in networks of all sizes
- Avoids aggregation bias
- Correctly handles non-collapsible effect measures
- Produces estimates in any target population for decision making
- Extends the standard network meta-analysis (NMA) framework, reducing to:
  - IPD network meta-regression with full IPD
  - Standard NMA with no adjustment
- Allows assumptions to be tested/relaxed in larger networks (Phillippo et al. 2023)
- Implemented in R package *multinma*

ML-NMR

1. Define an individual-level regression model
    - IPD network meta-regression, the gold-standard approach
    - Survival functions of any form: parametric PH or AFT, splines or piecewise exponential etc.
2. Average (integrate) the likelihood over the aggregate populations to form the aggregate-level model
    - Use general numerical integration



Predicting quantities of interest for a target population

The target population could be represented by

- A randomised trial
- An observational study
- A registry dataset
- …

With IPD covariate information

1. Make predictions for each individual
2. Summarise these for the population

With summary statistics

1. Generate integration points from joint covariate distribution
2. Integrate over the target population

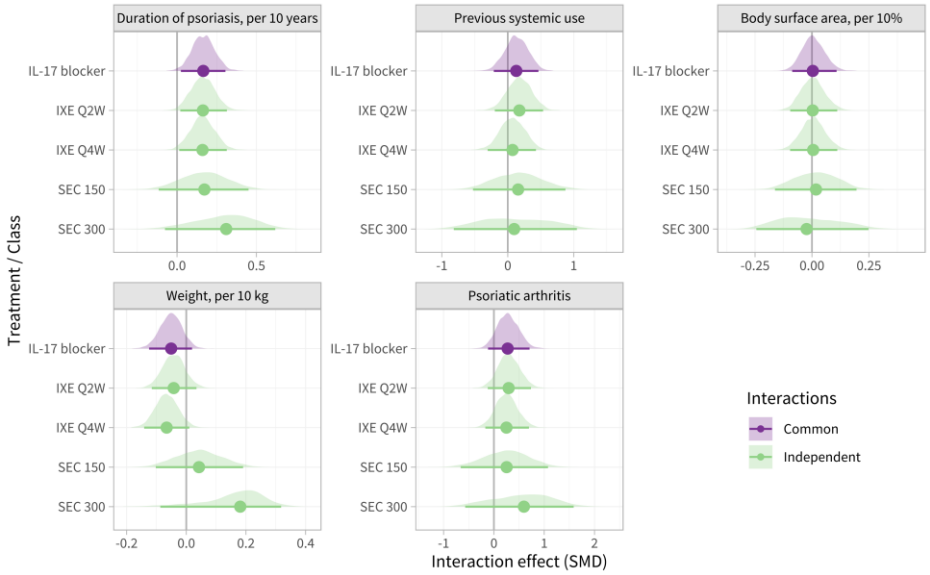
Absolute/marginal predictions need target population baseline hazard

- Modelled from available KM data or borrowed from a trial in the network

Assessing assumptions with ML-NMR

- Violation of conditional constancy (e.g. unobserved effect modifiers) may be detected using standard NMA methods
  - Random effects models – residual heterogeneity
  - Inconsistency models – residual inconsistency

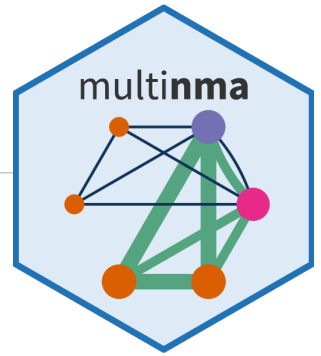
- Shared effect modifier assumption may be relaxed, one covariate at a time in smaller networks
  - Often needed to identify interaction terms for AgD treatments
  - If not valid then estimates for these treatments cannot be transported (like MAIC and STC)



Phillippo et al. (2023)

University of BRISTOL

The multinma R package



A user-friendly and comprehensive suite of tools for performing NMA and ML-NMR with AgD, IPD, or mixtures of both

- Models fitted in a Bayesian framework using Stan

Outcomes, likelihoods, link functions:

- **Binary/count** –Bernoulli/binomial (logit, probit, cloglog)
- **Rate** – Poisson (log)
- **Ordered categorical** – Multinomial (logit, probit, cloglog)
- **Continuous** – Normal (identity, log)
- **Survival** – Exponential (PH/AFT), Weibull (PH/AFT), Gompertz, log-Normal, log-Logistic, Gamma, Generalised Gamma, M-spline, piecewise exponential
- Includes features for modelling non-proportional hazards

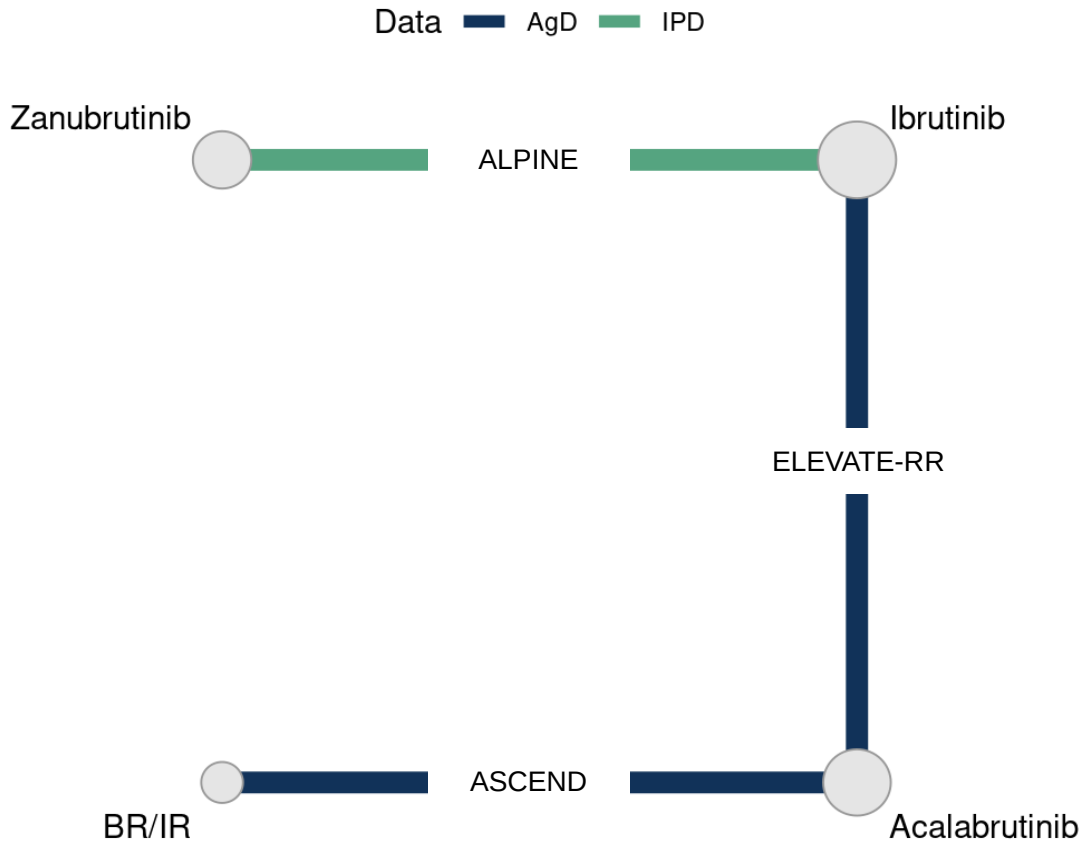
Case study: relapsed/refractory chronic lymphocytic leukemia

BeiGene ML-NMR analysis

- Four treatments:
  - Zanubrutinib
  - Ibrutinib
  - Acalabrutinib

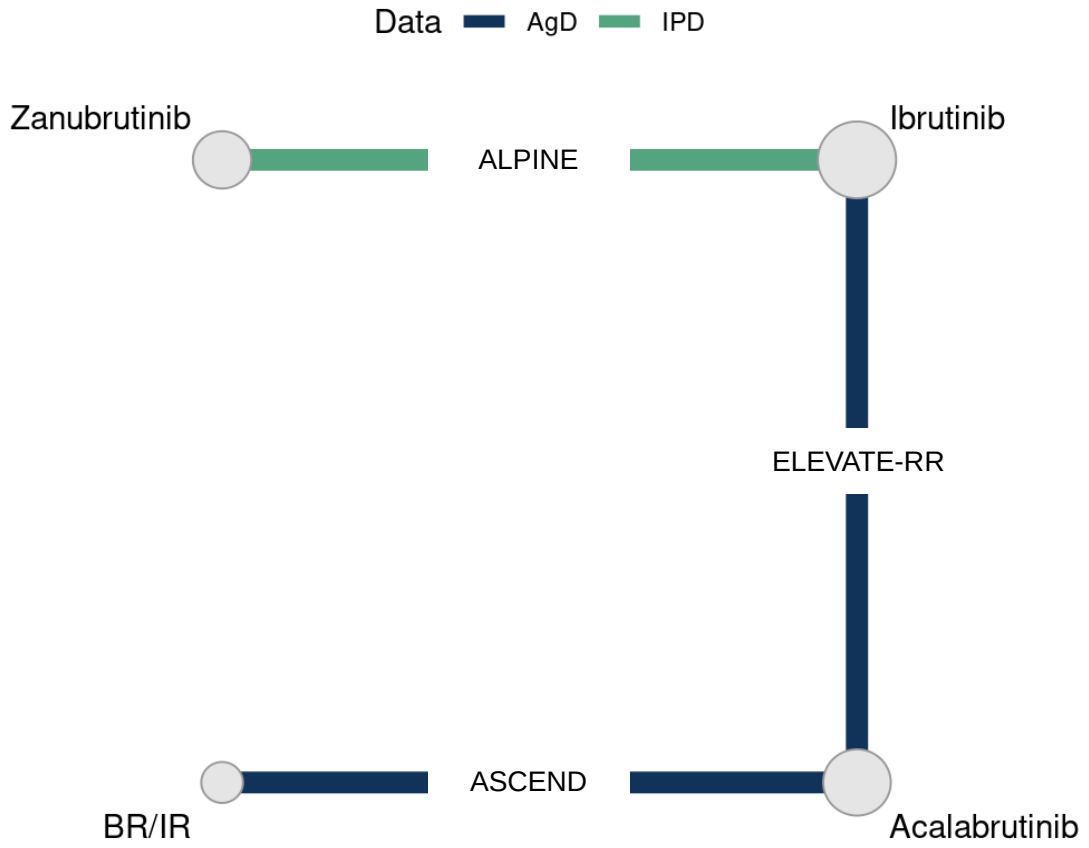
  } Bruton Tyrosine Kinase (BTK) inhibitors

  - Idelalisib plus rituximab (IR) or bendamustine plus rituximab (BR)
- IPD for ALPINE
- AgD for ELEVATE-RR and ASCEND
- PFS and OS outcomes
- MAIC and STC not possible

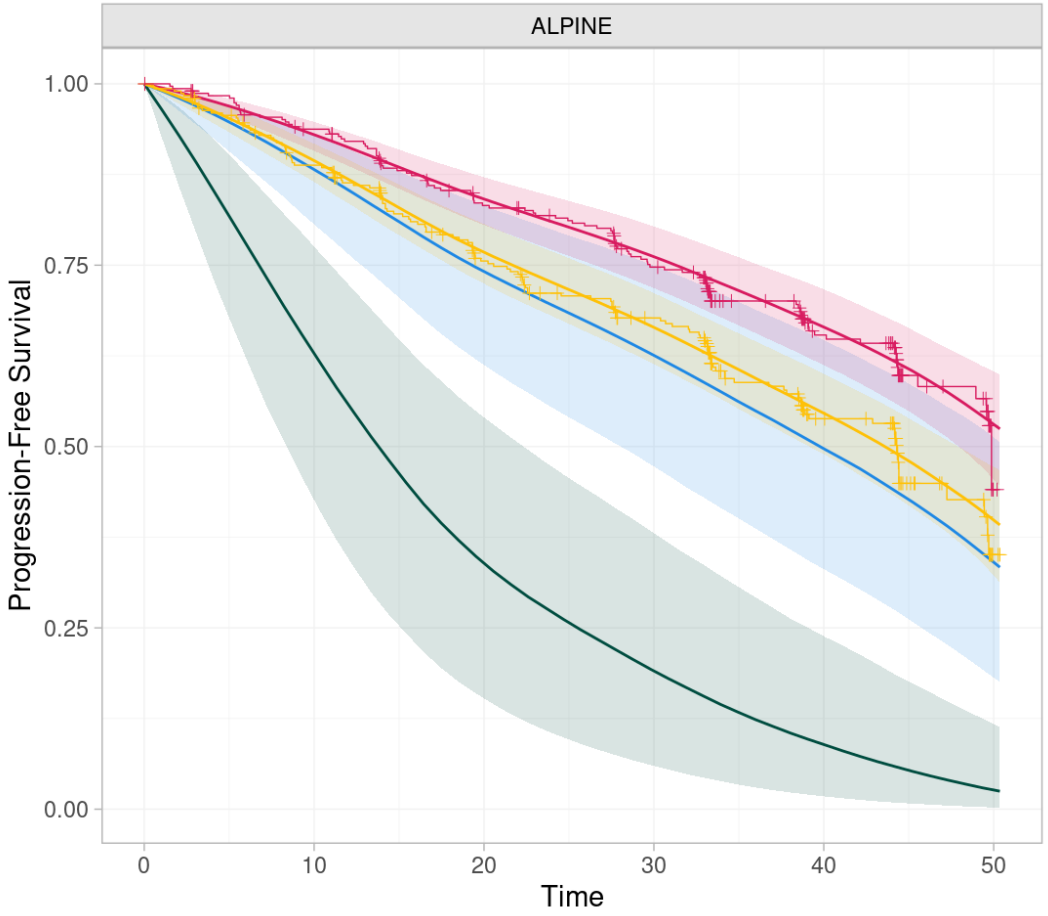


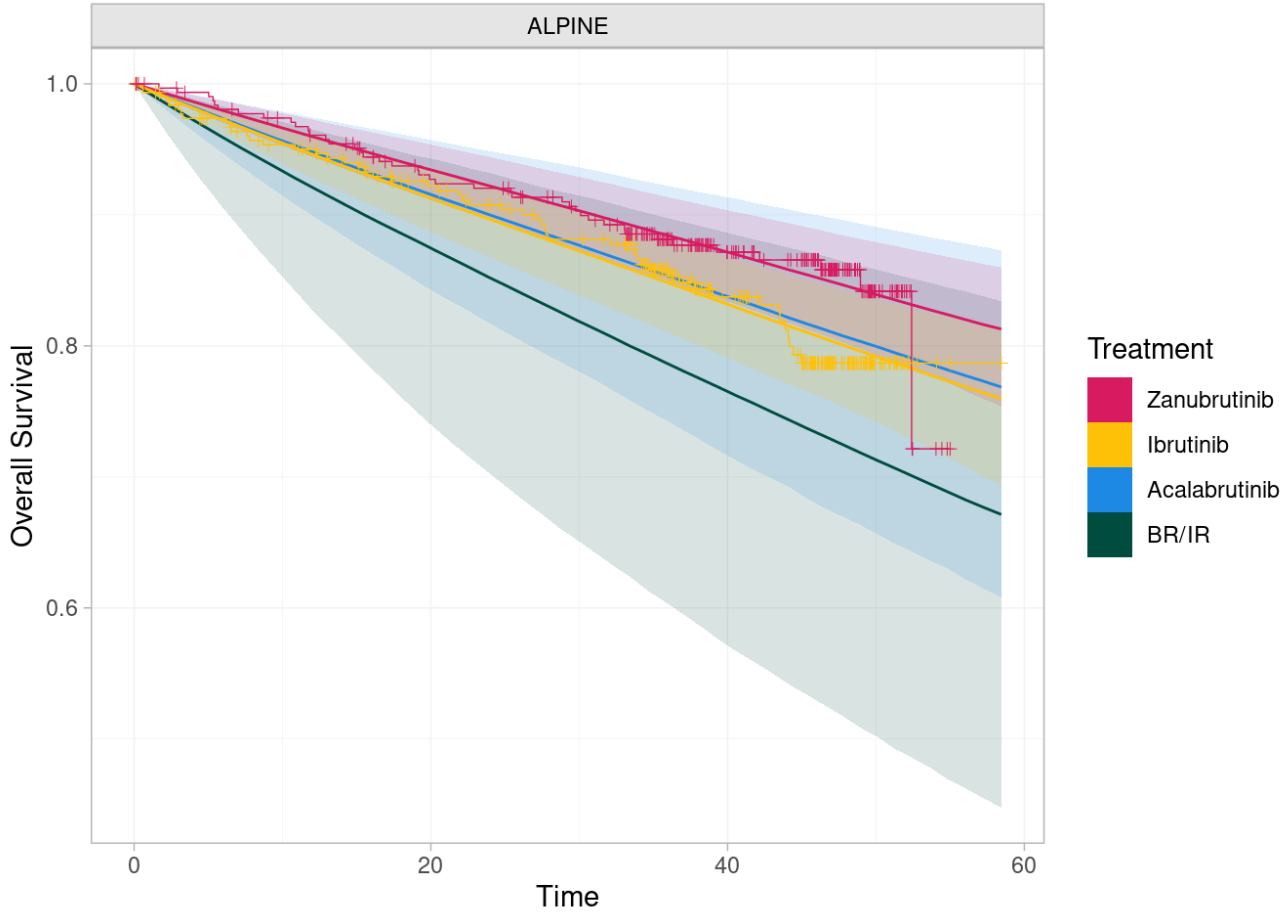
Case study: relapsed/refractory chronic lymphocytic leukemia

- 2 target populations:
  - ALPINE
  - High-risk subgroup del(17p) / del(11q)
- 7 potential effect modifiers:
  - Age
  - Geographic region (Asia versus rest)
  - Rai/Binet stage (0-II vs. III-IV)
  - Bulky disease (≥5cm)
  - del(17p), TP53, and del(11q) mutations
  - +5 more for sensitivity analyses
- Shared EM assumption between zanubrutinib, acalabrutinib, and BR/IR
- Fitted parametric and flexible models



Case study: results – ALPINE PFS and OS





- M-splines selected as best-fitting model. Good visual fit, M-splines capture shape of baseline hazard well
- All three BTK inhibitors better than BR/IR for both PFS and OS

University of BRISTOL

Case study: results – population-average conditional HRs

HR for PFS	ALPINE ITT population	High-risk population
Zanubrutinib vs Acalabrutinib	0.57 (0.34, 0.95)	0.57 (0.34, 0.95)
Zanubrutinib vs Ibrutinib	0.67 (0.52, 0.87)	0.47 (0.29, 0.73)
Zanubrutinib vs BR/IR	0.15 (0.08, 0.28)	0.15 (0.08, 0.28)

HR for OS	ALPINE ITT population	High-risk population
Zanubrutinib vs Acalabrutinib	0.77 (0.36, 1.65)	0.77 (0.36, 1.65)
Zanubrutinib vs Ibrutinib	0.67 (0.40, 1.07)	0.58 (0.30, 1.06)
Zanubrutinib vs BR/IR	0.48 (0.19, 1.19)	0.48 (0.19, 1.19)

- Zanubrutinib estimated to improve PFS and OS against other treatments, in both ALPINE and high-risk populations; 95% credible intervals exclude 1 for PFS, more uncertainty for OS
- All three BTK inhibitors better than physician's choice of BR/IR
- Align with results of previous unadjusted network meta-analysis
- Notice that zanubrutinib – acalabrutinib – BR/IR comparisons are constant across populations

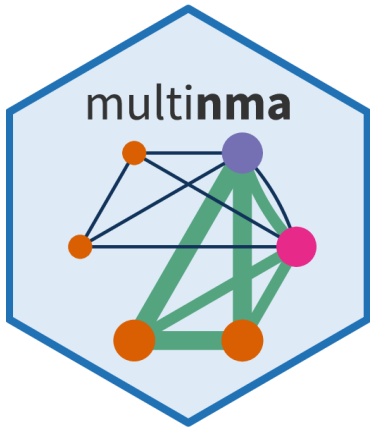
Case study: advantages and limitations

- Adjusted for differences in effect-modifying covariates between studies
- EMs identified by subgroup analyses, systematic review, clinical judgement
- Produced estimates for two target populations of interest
- Robust to results of sensitivity analyses with different covariates
- MAIC and STC analyses not possible, relaxed assumptions vs previous NMA

- Shared EM assumption required between zanubrutinib, acalabrutinib, and BR/IR
  - BR/IR vs acalabrutinib / zanubrutinib comparisons (incl ASCEND) assumed constant, unaltered by EMs
  - May not be a feasible assumption for BR/IR, only affects estimates and comparisons with BR/IR
  - More likely to be reasonable for next-gen BTK inhibitors
- Small network, cannot assess unobserved EMs with heterogeneity/inconsistency checks

Summary

- ML-NMR is a flexible and general method for synthesising evidence from mixtures of individual and aggregate level data
- Several advantages over previous population-adjustment methods
  - Coherently analyse networks of any size
  - Produce estimates in a relevant decision target population
  - Assess key assumptions in larger networks
- Uptake and acceptance in NICE appraisals (TA912, TA1013)
- Implemented in multinma R package
  - Website: dmphillippo.github.io/multinma
  - Documentation, example analyses



Thank you

david.phillippo@bristol.ac.uk

References

Shadman, M et al. (2024) Zanubrutinib vs other Bruton tyrosine kinase inhibitors in relapsed/refractory chronic lymphocytic leukemia: a multilevel network meta-regression. Poster presented at *European Hematology Association Congress 2024,* Madrid, Spain.

Phillippo, DM et al. (2020) Multilevel Network Meta-Regression for population-adjusted treatment comparisons. *Journal of the Royal Statistical Society: Series A*, 183(3):1189-1210. DOI: 10.1111/rssa.12579.

Phillippo, DM et al. (2020) Assessing the performance of population adjustment methods for anchored indirect comparisons: A simulation study. *Statistics in Medicine*, 39(30):4885-4911. DOI: 10.1002/sim.8759

Phillippo DM et al. (2023) Validating the assumptions of population adjustment: application of multilevel network meta-regression to a network of treatments for plaque psoriasis. *Medical Decision Making*, 43(1):53-67. DOI: 10.1177/0272989X221117162.

Phillippo DM et al. (2024) Multilevel network meta-regression for general likelihoods: synthesis of individual and aggregate data with applications to survival analysis. Preprint, *arXiv*:2401.12640 [stat.ME]

R package *multinma*, see dmphillippo.github.io/multinma for details