METRUM
RESEARCH GROUP

Graybill Conference
Fort Collins, CO

Evaluating Conditional Exchangeability Assumptions for Bayesian Borrowing, With Application to Pediatric Extrapolation

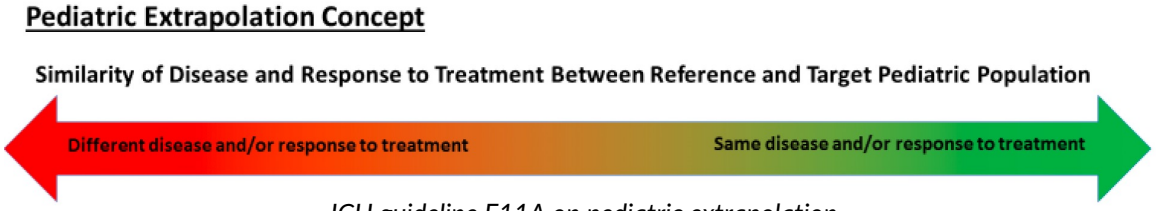
June 12, 2024

James A. Rogers, Ph.D. (Metrum RG)
Curtis Johnston, Pharm.D. (Metrum RG)
Matthew Wiens, Ph.D. (Metrum RG)
Alejandro Perez-Pitarch, Ph.D. (Boehringer Ingelheim*)
Hugo Maas (Boehringer Ingelheim)

** affiliation when work was done; APP is now employed by Regeneron*

The Consulting Challenge

How do we get from this world:



ICH guideline E11A on pediatric extrapolation

To this world:

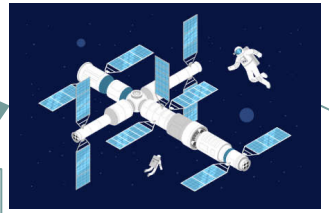
The structure of the problem permits us to satisfy condition 2 of Theorem 3, since Z is S-admissible and $P^*(z|\,\mathrm{do}(x))$ is trivially transportable. The former can be seen from $(S \perp\!\!\!\perp Y|X, Z)_{G_{\overline{X}}}$, hence $P^*(y|\,\mathrm{do}(x), z) = P(y|\,\mathrm{do}(x), z)$; the latter can be seen from the fact that X and Z and unconfounded, hence $P^*(z|\,\mathrm{do}(x)) = P^*(z|x)$. Putting the two together, we get

$$P^*(y|\,\mathrm{do}(x)) = \sum_z P(y|\,\mathrm{do}(x), z)\,P^*(z|x), \quad (5.8)$$

Pearl and Bareinboim, External Validity: From Do-Calculus to Transportability Across Populations, Stat. Sci. 2014.

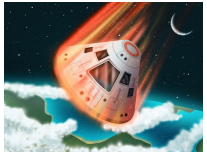
And back again?

Mathematical deductions



Isolated / “atomic” statements about formal exchangeability (DAGs, selection diagrams)

Abstract statement of consequences



Elicitation

Interpretation

Isolated / “atomic” qualitative statements about similarity

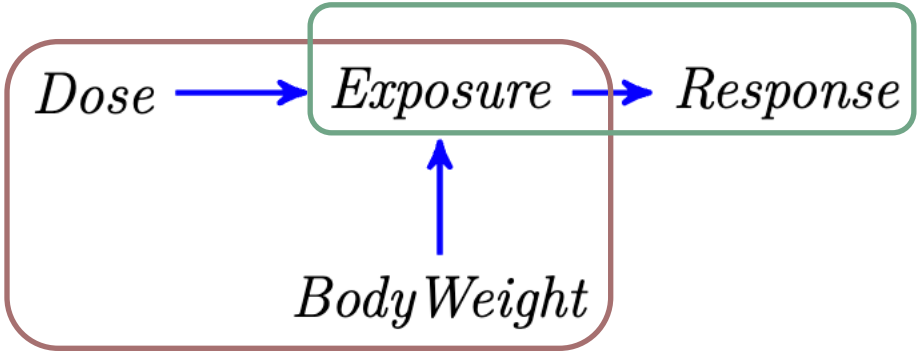
Claim about validity of extrapolation plan as a whole

Mutual / bidirectional critique / cross-checking

linguistic arguments

Causal DAGs as Summaries of Within-Group Similarity Statements

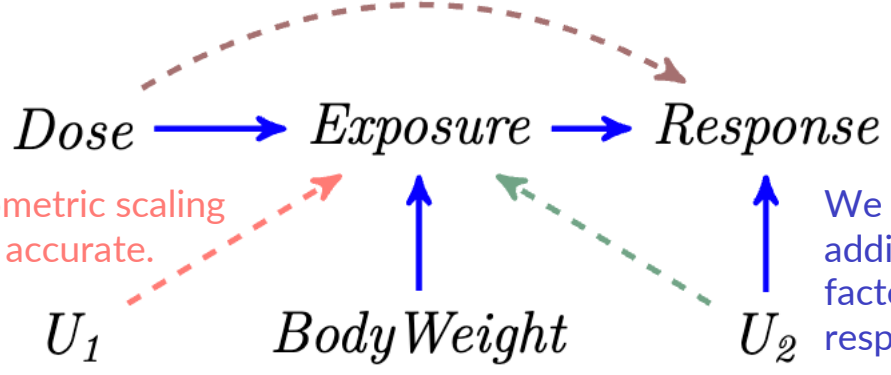
"We expect 2 different adults to have similar responses if they have similar exposure (even if they have different doses and/or bodyweights)"



"We expect 2 different adults to have similar exposure if they have the same dose and similar body weight"

Dialogue About Similarity Assumptions

Based on E-R analyses, substantial variation in the response is explained by variation in our measure of exposure, so we assume complete mediation



PK model with just allometric scaling by body weight is very accurate.

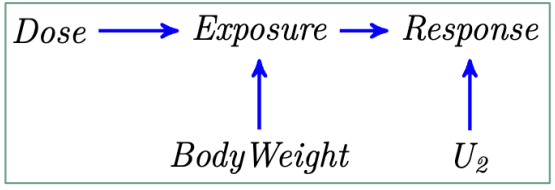
We do expect additional unknown factors to affect PD response …

… but there is no easily-hypothesized pharmacology to suggest overlap in PK and PD factors, so we assume no confounding in E-R.

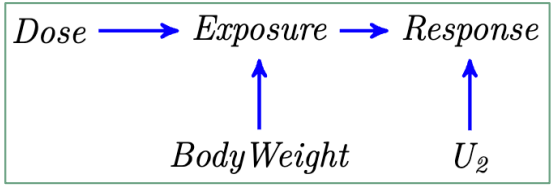
*NB: these are just **example** justifications that could be given for excluding the dashed arrows, **depending on context***

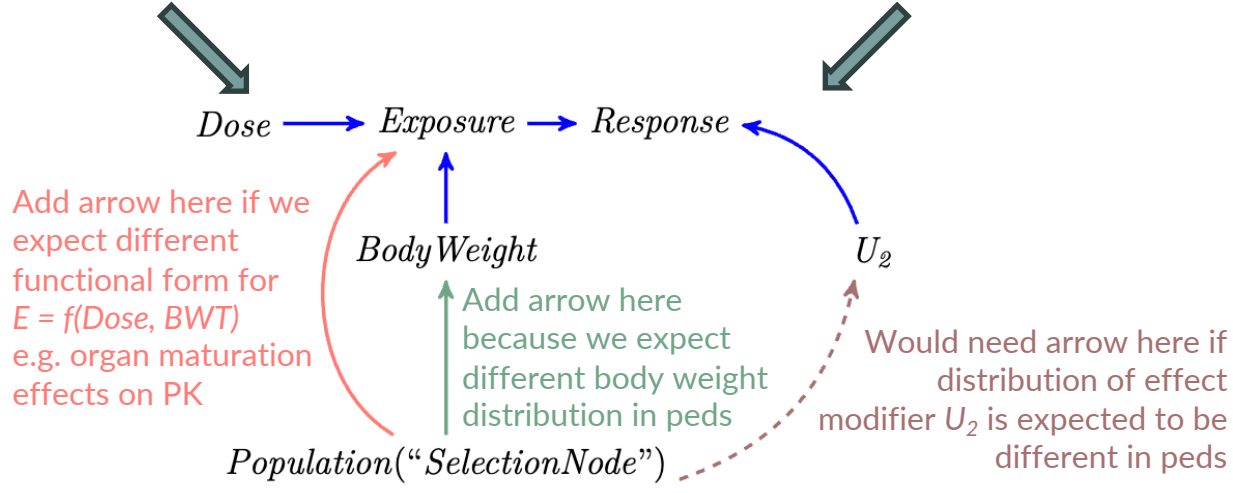
Between-Group (Dis)similarities

Working assumptions regarding similarities *within* adult population



Working assumptions regarding similarities *within* pediatric population





(this is what Pearl and Bareinboim call a "selection diagram")

The Four Probability Spaces of Pediatric Extrapolation

Let $\boldsymbol{v}$ represent all variables in the within-group DAGs

We may not need this, if we have the right randomized studies in adults

We usually know some features of this distribution from a variety of sources, e.g. epi databases like NHANES, early phase ped data

	Adult	Pediatric
Observational	$P(\boldsymbol{v}\|D=d)$	$P^*(\boldsymbol{v}\|D=d)$
Interventional	$P(\boldsymbol{v}\|\text{do}(D=d))$	$P^*(\boldsymbol{v}\|\text{do}(D=d))$

We know some features of this distribution, if we have randomized studies in adults

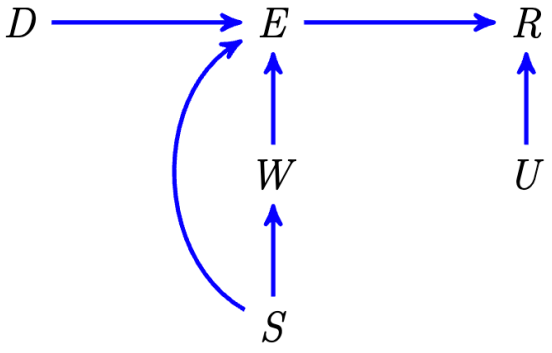
We likely need randomized studies in pediatrics to learn some features of this distribution, *but not necessarily studies of the clinical endpoint*

Goal: cobble these pieces together to determine $P^*(r|do(D=d))$

In Pearl and Bareinboim's terminology, a formula that accomplishes this is a "transport formula"

METRUM RESEARCH GROUP

Selection Graph & Transport Formula for "Full Extrapolation" / "Exposure Matching"



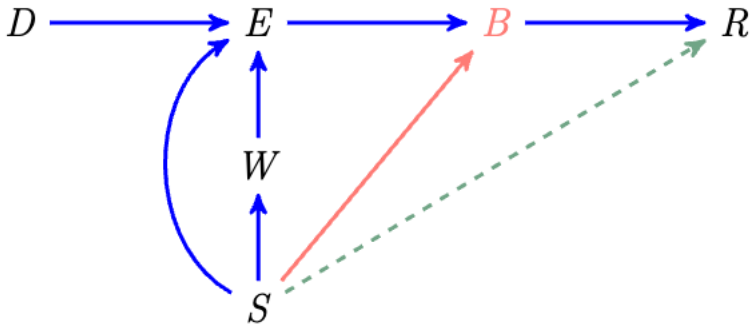
D = dose or treatment status
E = exposure (summary metric)
R = Response
W = Body weight
U = Unmeasured effect modifier
S = Selection node (pediatric status)

Transport formula:

$$
\begin{aligned}
P^*(r \mid \mathrm{do}(D=d)) &= \int_e P^*(r \mid e, \mathrm{do}(D=d)) P^*(e \mid \mathrm{do}(D=d)) \mathrm{d}e && \text{(Law of total prb.)} \\
&= \int_e P^*(r \mid e) P^*(e \mid \mathrm{do}(D=d)) \mathrm{d}e && (D \perp R \mid E) \\
&= \int_e \underbrace{P(r \mid e)}_{\text{Estimate with E-R model fit to adult data}} \underbrace{P^*(e \mid \mathrm{do}(D=d))}_{\text{Estimate with randomized study in peds with exposure endpoint}} \mathrm{d}e && \underbrace{(S \perp R \mid E)}_{\text{Conditional exchangeabilities derived from selection diagram}}
\end{aligned}
$$

For example of justification to support full extrapolation, see Kalaria et al. CPT 2019.

Selection Graph & Transport Formula for "Bridging Biomarker" Approach



Introducing bridging biomarker *B* may make it easier to justify removal of *S* → *R*.

Especially true if *B* is known to be causally proximate to *R*

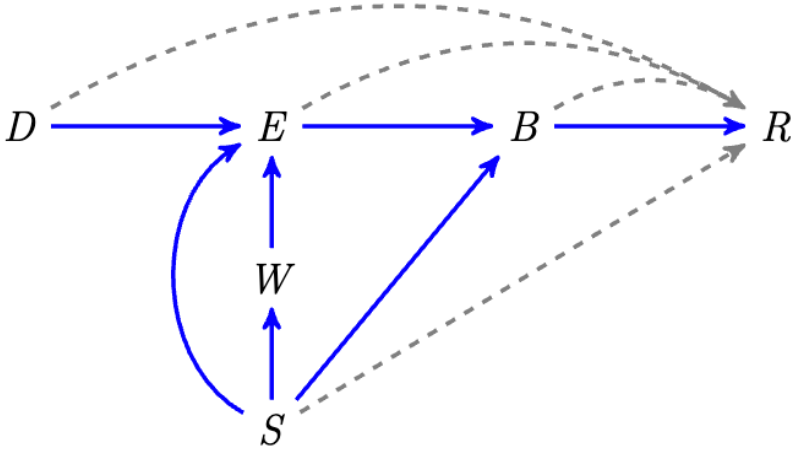
When this selection diagram does not include *S* → *R* :

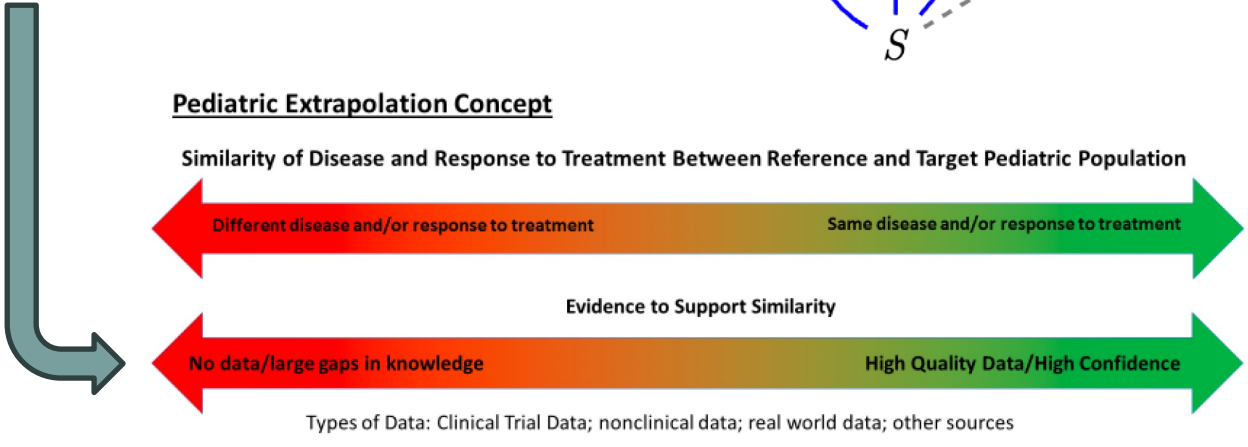
$$
\begin{aligned}
P^*(r \mid \mathrm{do}(D=d)) &= \textstyle\int_b P^*(r \mid b, \mathrm{do}(D=d)) P^*(b \mid \mathrm{do}(D=d)) \mathrm{d}b && \text{(Law of total prb.)} \\
&= \textstyle\int_b P^*(r \mid b) P^*(b \mid do(D=d)) \mathrm{d}b && (D \perp R \mid B) \\
&= \textstyle\int_b \underbrace{P(r \mid b)}_{\text{Estimate with disease progression model fit to adult data}} \underbrace{P^*(b \mid do(D=d))}_{\text{Estimate with randomized study in peds with biomarker endpoint}} \mathrm{d}b && (S \perp R \mid B)
\end{aligned}
$$

For discussion of bridging biomarkers, see Fleming et al. Ther. Innov. & Reg. Sci. 2022

What About All Those Arrows We Deleted?

- Diagram creation fosters good conversations about assumptions (this is already a win)
- But: selection diagrams either include arrows or they don't
- If used in isolation, there is no room for the continuum of evidence described in ICH E11a



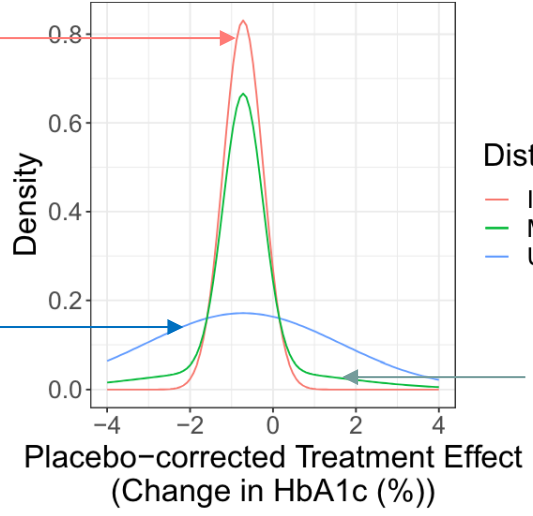


Bayesian Approach to Respect Continuum in Strength of Prior Evidence

- Bayesian prior based on working hypothesis that similarity assumptions are correct
- Robustify prior to acknowledge that assumptions / selection diagram could be wrong

Maybe our conditional exchangeability assumptions are correct …

… and maybe they aren't



Working prior gives *partial* credence to conditional exchangeability assumptions

Sailer, O., et al. Pharmacometrics-enhanced Bayesian borrowing for paediatric extrapolation - A case study of the DINAMO trial. PSI London (2023).
Johnston, C., et al. Bayesian Borrowing in the DINAMO Pediatric Study using Informative Priors Derived from Model-based Extrapolation. American Conference on Pharmacometrics (2023).

What Do We Gain With Diagrams?

01 If we take pains to develop a fancy selection diagram encoding conditional exchangeability assumptions,

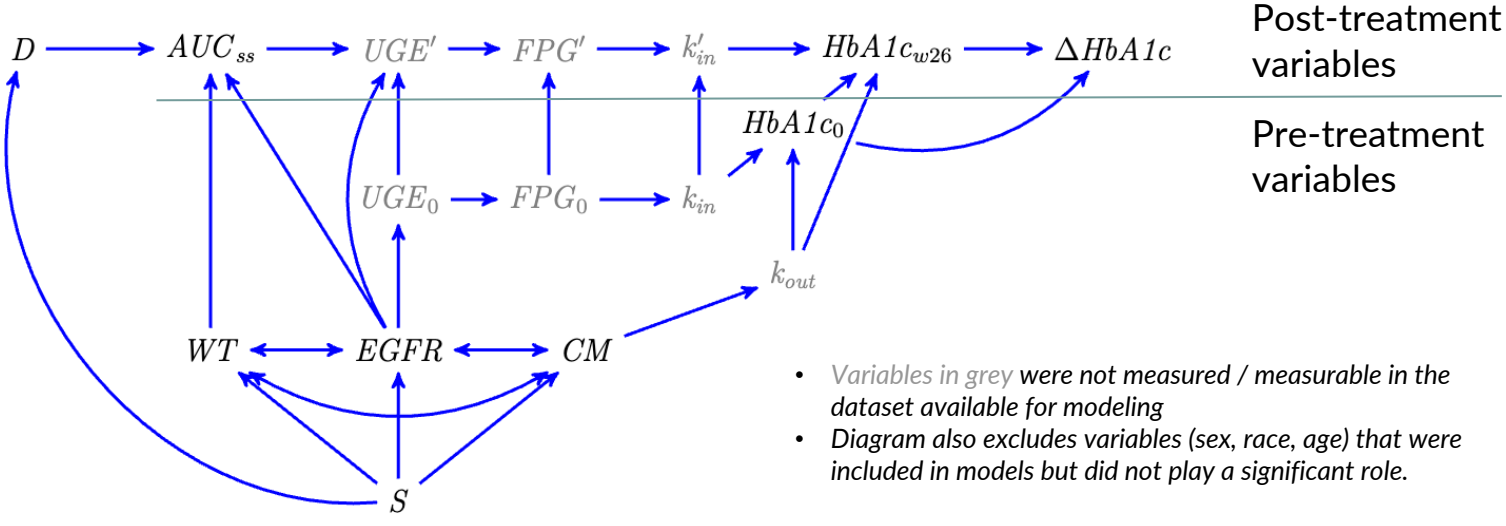
02 We still seem to give up:
"Maybe it's right, maybe it isn't; let's just be Bayesian"

03 However: along the way, **a richer conversation** about what we believe and why we believe it
(a consulting victory, not a Q.E.D.)

04 Result: more **transparent and collaborative justification of prior** + **better planning** to eliminate evidence gaps

"Among biostatisticians working in later phase drug trials, the working group observes that reluctance to use Bayesian methods appears to have three primary causes. First, the Bayesian approach does require an initial assessment of the commensurability of the various sources of information, which is often difficult for investigators to make."
Gamalo-Siebers et al, Statistical modeling for Bayesian extrapolation of adult clinical trial information in pediatric drug evaluation. Pharm Stat. 2017

Selection Diagram For SGLT-2 Inhibitor Partial Extrapolation

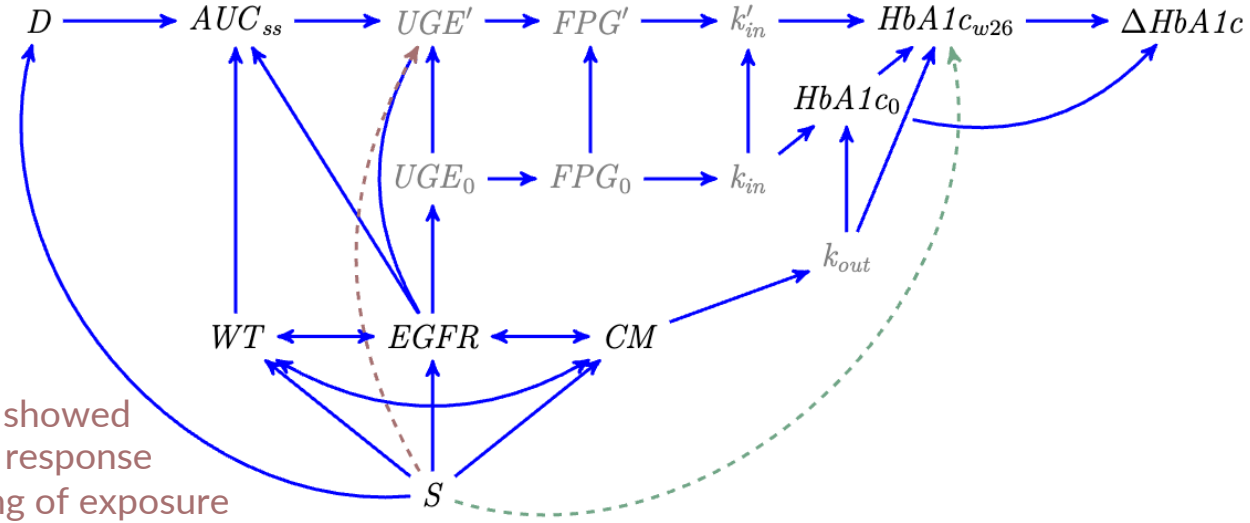


- *Variables in grey were not measured / measurable in the dataset available for modeling*
- *Diagram also excludes variables (sex, race, age) that were included in models but did not play a significant role.*

D = dose or treatment status
AUC_{ss} = exposure at PK steady-state
UGE_0, UGE' = urinary glucose excretion, pre- and post-treatment
FPG_0, FPG' = fasting plasma glucose, pre- and post-treatment
k_{in}, k'_{in} = rate constant for HbA1c synthesis, pre- and post-treatment
k_{out} = rate constant for HbA1c degradation / elimination

$HbA1c_0, HbA1c_{w26}$ = hemoglobin A1c at baseline and week 26
$\Delta HbA1c$ = Change from baseline in hemoglobin A1c
WT = Body weight
$EGFR$ = estimated glomerular filtration rate
CM = Concomitant medication usage
S = Selection node (pediatric status)

Evidence of Similarity in Response to Treatment

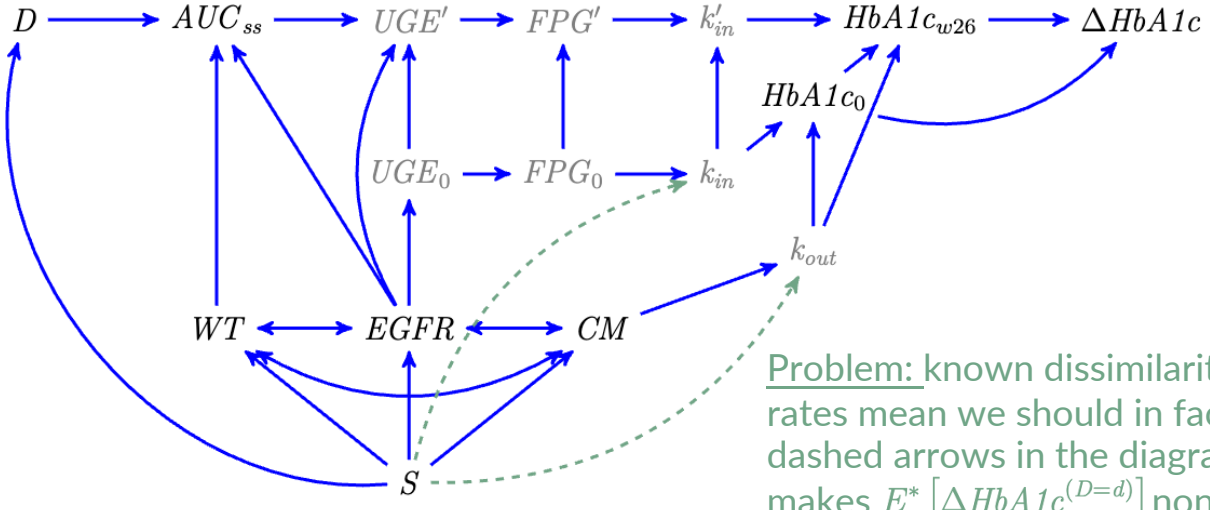


Prior study of *UGE* in pediatrics and adults showed no difference in *UGE* response after allometric scaling of exposure and adjustment for *EGFR*. Good evidence to remove the *S* → *UGE'* arrow

Known pharmacology and relationships between *UGE*, fasting plasma glucose, and *HbA1c* (*) imply most or all of the drug effect is mediated through *UGE*. Justifies removal of *S* → $HbA1c_{w26}$ arrow, which in turn removes this threat to identifiability.

** Described in Riggs et al. Exposure–response modelling for empagliflozin, a sodium glucose cotransporter 2 (SGLT2) inhibitor, in patients with type 2 diabetes, BJCP 2014.*

Robustness to Dissimilarity in Disease Progression



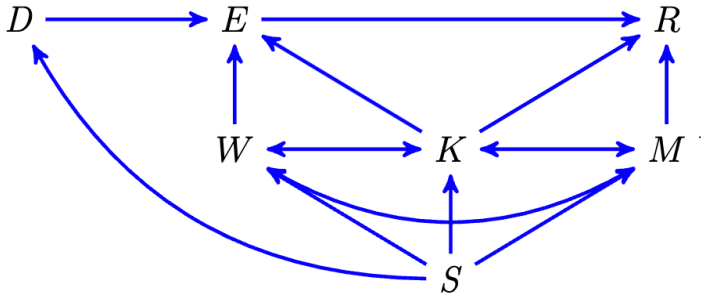
Problem: known dissimilarities in progression rates mean we should in fact include these dashed arrows in the diagram, but doing so makes $E^*\left[\Delta HbA1c^{(D=d)}\right]$ non-identifiable

Solution: compute (by simulation) $E^*\left[\Delta HbA1c^{(D=d)} - \Delta HbA1c^{(D=0)}\right]$ *as if* the arrows were absent, then show that this is conservative (pessimistic w.r.t. efficacy) relative to sensitivity analyses with hypothesized dissimilarities in k_{in} and k_{out} (fortunately, this turns out to be the case)

Simplifying Selection Diagrams to Derive Transport Formulas

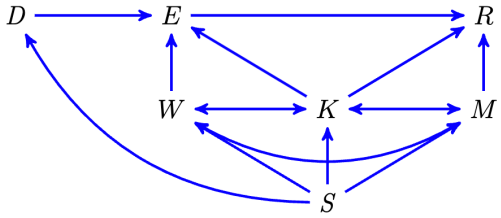
On previous slides we developed a credible selection diagram to represent expert beliefs

For purposes of checking a transport formula, we may **simplify the diagram without changing its mathematical implications** by removing unmeasured intermediates and variables that we don't need to condition on



D = dose or treatment status
E = exposure at PK steady-state
R = Change from baseline in hemoglobin A1c
W = Body weight
K = Kidney function (EGFR)
M = Concomitant medication usage

More Complex Transport Formula



$$
\begin{aligned}
&P^*(r \mid \mathrm{do}(d)) \\
&\quad = \int_w \int_k \int_m P^*(r \mid \mathrm{do}(d), w, k, m) P^*(w, k, m \mid \mathrm{do}(d)) \mathrm{d}m\mathrm{d}k\mathrm{d}w && \text{(Law of t.p.)} \\
&\quad = \int_w \int_k \int_m P^*(r \mid \mathrm{do}(d), w, k, m) P^*(w, k, m) \mathrm{d}m\mathrm{d}k\mathrm{d}w && ((W, K, M) \perp D)_{G_{\overline{D}}}
\end{aligned}
$$

$$
\begin{aligned}
&= \int_e P^*(r \mid \mathrm{do}(d), e, w, k, m) P^*(e \mid w, k, m, \mathrm{do}(d)) \mathrm{d}e && \text{(Law of t.p.)} \\
&= \int_e P(r \mid \mathrm{do}(d), e, w, k, m) P^*(e \mid w, k, m, \mathrm{do}(d)) \mathrm{d}e && (R \perp S \mid W, K, M)_{G_{\overline{D}}} \\
&= \int_e P(r \mid \mathrm{do}(d), e, w, k, m) P^*(e \mid w, k, \mathrm{do}(d)) \mathrm{d}e && (E \perp M \mid W, K)_{G_{\overline{D}}} \\
&= \int_e P(r \mid e, k, m) P^*(e \mid w, k, \mathrm{do}(d)) \mathrm{d}e && (R \perp (W, D) \mid W, K, M, E)_{G_{\overline{D}}}
\end{aligned}
$$

$$
P^*(r \mid \mathrm{do}(d)) = \int_w \int_k \int_m \int_e P(r \mid e, k, m) P^*(e \mid w, k) P^*(w, k, m) \mathrm{d}m\mathrm{d}k\mathrm{d}w\mathrm{d}e
$$

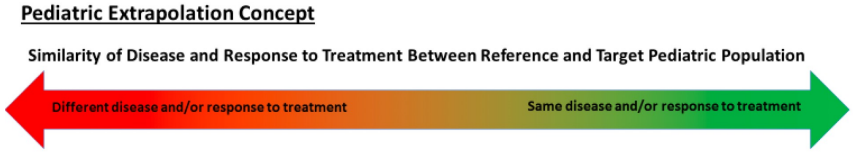
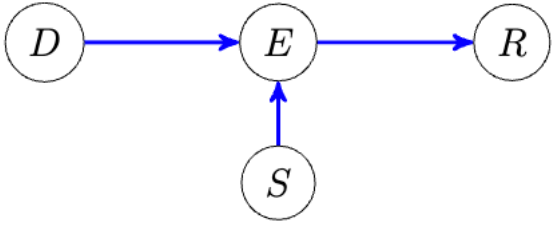
Estimate with E-R or PKPD model fit to adult data

Estimate with PK model fit to pediatric (or ped + adult) data

Estimate with empirical pediatric dist'n or epi database

In practice, the above integrals are estimated by averaging over Monte-Carlo simulations from the outcome models

Selection Diagrams Bridge Between Worlds

Non-statistician mood	Summary of assumptions	Statistician mood
		
		
	The structure of the problem permits us to satisfy condition 2 of Theorem 3, since Z is S-admissible and $P^*(z\|\,\mathrm{do}(x))$ is trivially transportable. The former can be seen from $(S \perp\!\!\!\perp Y \| X, Z)_{G_{\underline{X}}}$, hence $P^*(y\|\,\mathrm{do}(x), z) = P(y\|\,\mathrm{do}(x), z)$; the latter can be seen from the fact that X and Z and unconfounded, hence $P^*(z\|\,\mathrm{do}(x)) = P^*(z\|x)$. Putting the two together, we get (5.8) $P^*(y\|\,\mathrm{do}(x)) = \sum_z P(y\|\,\mathrm{do}(x), z)\, P^*(z\|x)$,	

Thank You

Get in touch with me:

jimr@metrumrg.com

References

Pearl, J. and Bareinboim, E. External Validity: From Do-Calculus to Transportability Across Populations. Stat Sci. 2014, Vol. 29, No. 4, 579–595

EMA. Draft ICH guideline E11A on pediatric extrapolation Step 2b. 2022.

Kalaria, S. et al. Extrapolation of Efficacy and Dose Selection in Pediatrics : A Case Example of Atypical Antipsychotics in Adolescents With Schizophrenia and Bipolar I Disorder. 2021, Vol. 61, No. S1, S117-S124.

Fleming, T. et al. Innovations in Pediatric Therapeutics Development: Principles for the Use of Bridging Biomarkers in Pediatric Extrapolation. Ther. Innov. & Reg. Sci. 2022, Vol. 57, 109-120.

Sailer, O., et al. Pharmacometrics-enhanced Bayesian borrowing for paediatric extrapolation - A case study of the DINAMO trial. PSI London (2023).

Johnston, C., et al. Bayesian Borrowing in the DINAMO Pediatric Study using Informative Priors Derived from Model-based Extrapolation. American Conference on Pharmacometrics (2023).

Gamalo-Siebers, M. et al. Statistical modeling for Bayesian extrapolation of adult clinical trial information in pediatric drug evaluation. Pharm Stat. 2017, Vol. 16, Issue 4, 232-249.

Ye, J. et al. Recent Use of Pediatric Extrapolation in Pediatric Drug Development in US. J. Biopharm. Stat. 2023, Vol. 33, No. 6, 681-695.

Fayette, L., Sailer, O., and Perez-Pitarch, A. Pharmacometrics enhanced Bayesian borrowing approach to improve clinical trial efficiency: Case of empagliflozin in type 2 diabetes. CPT-PSP. 2023, Vol. 12, No. 10, 1386-1397.

Riggs, M. et al. Exposure–response modelling for empagliflozin, a sodium glucose cotransporter 2 (SGLT2) inhibitor, in patients with type 2 diabetes. BJCP. 2014. Vol. 78, No. 6, 1407-1418.

> *Science is about generalization, and generalization requires that conclusions obtained in the laboratory be transported and applied elsewhere, in an environment that differs in many aspects from that of the laboratory... On the theoretical front, the standard literature on [extrapolation], falling under rubrics such as "external validity" …consists primarily of "threats," namely, explanations of what may go wrong when we try to transport results from one study to another while ignoring their differences. … this paper departs from the tradition of communicating "threats" and embarks instead on the task of formulating "licenses to transport," namely, assumptions that, if they held true, would permit us to transport results across studies.*

— Judea Pearl and Elias Bareinboim, Stat Sci 2014