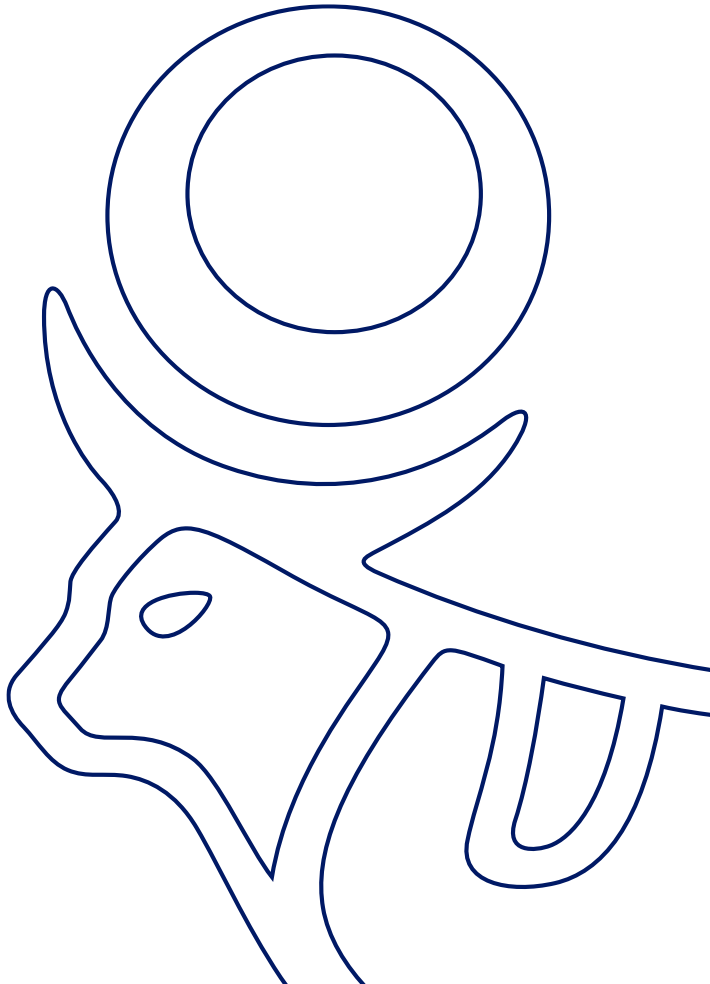
UNORDINARY DRIVES CHANGE

novo nordisk®

What is your estimand?

Navigating marginal and conditional treatment effects



Antonio Remiro-Azócar, PhD

Biostatistics Methods, Medical and Translational Science, Novo Nordisk

ISPOR Europe, Glasgow, 12th November 2025

Agenda

1.) Within-trial estimands

- Marginal versus conditional summary measures (“causal contrasts”)

2.) Implications for transportability and compatibility across studies

- Transportability: “purely prognostic” variables may modify marginal treatment effects
- Compatibility: trials may employ different analysis methods and report different summary measures in publications…when are summary measures compatible?

3.) Implications for evidence synthesis methodology

- Aggregate-level data meta-analysis
- Population-adjusted indirect treatment comparisons

Note: my focus is on comparative efficacy/effectiveness (pre-requisite for cost-effectiveness!)

Within-trial estimands

Marginal treatment effect

$$\text{MTE} = g\left(E\left(Y^1\right)\right) - g\left(E\left(Y^0\right)\right)$$

- A contrast of functions of marginal or unconditional outcome expectations under each treatment
- Typically described as the average treatment effect in the trial population…had everyone in the trial population been under the active treatment versus the control
- Of primary interest for population-level policy decisions based on clinical efficacy or effectiveness
- Sometimes referred to as the ATE, SATE or PATE
- Within-trial estimators: crude unadjusted difference in outcome means, simple univariable regression of outcome on treatment, model-based standardization, inverse probability of treatment weighting

Conditional average treatment effect

$$\text{CATE} = g\left(E\left(Y^1 \mid X = x\right)\right) - g\left(E\left(Y^0 \mid X = x\right)\right)$$

- Covariate-specific treatment effect that conditions on a specific value of the covariate(s)…the average treatment effect had a subset of individuals with a given covariate profile been assigned active treatment versus control
- For binary/categorical covariates, a subgroup- or subset-specific treatment effect
- Can approximate an individualized or subject-specific effect, but only if multiple relevant “effect modifiers” are conditioned on to account for treatment effect heterogeneity
- Of little interest for population-level decision-making based on clinical efficacy or effectiveness, but can be useful when considering decisions for different subgroups if the treatment effect varies between subjects
- Within-trial estimators: direct regression adjustment, but with appropriate treatment-covariate interaction terms

Conditional treatment effect at the (covariate) means

$$\text{CTEM} = g\left(E\left(Y^1 \mid X = \bar{X}\right)\right) - g\left(E\left(Y^0 \mid X = \bar{X}\right)\right)$$

- Special case of the CATE: conditional average effect for subjects with covariate values equal to their mean
- Not meaningful and may have an awkward or confusing interpretation for discrete covariates
- Consider ethnicity, sex, smoking status or biomarker status, where the mean is a proportion of subjects in a category…the effect for a non-existent “average subject” does not make much sense
- Not a desirable target for population-level decision-making based on clinical efficacy or effectiveness
- Within-trial estimators: contrasts of “least squares means” (sometimes referred to as “marginal means”)

Population-average conditional treatment effect

$$\text{PACTE} = E_X\left(g\left(E\left(Y^1 \mid X = x\right)\right) - g\left(E\left(Y^0 \mid X = x\right)\right)\right)$$

- Is also an average treatment effect across the trial population…had everyone in the trial population been under active treatment versus control
- The average CATE across all subjects or subgroups in the trial population
- The difference between the MTE and the PACTE is the order of operations, which is irrelevant where the link function is the identity link but otherwise matters
- MTE: averages the (transformed) unconditional outcomes, then contrasts
- PACTE: contrasts the (transformed) conditional outcomes, then averages
- Within-trial estimators: direct regression adjustment, Mantel-Haenszel (averaging the stratum-specific CATEs)

Transportability and compatibility across studies

Model-based estimands

We postulate some hypothetical outcome-generating models to examine estimand behavior

Let’s consider three distinct parametric outcome-generating mechanisms:

1.) Homogeneous (constant) CATE across the covariate(s)

2.) Heterogeneous CATE varying linearly with the covariate(s)

3.) Heterogeneous CATE varying non-linearly with the covariate(s)

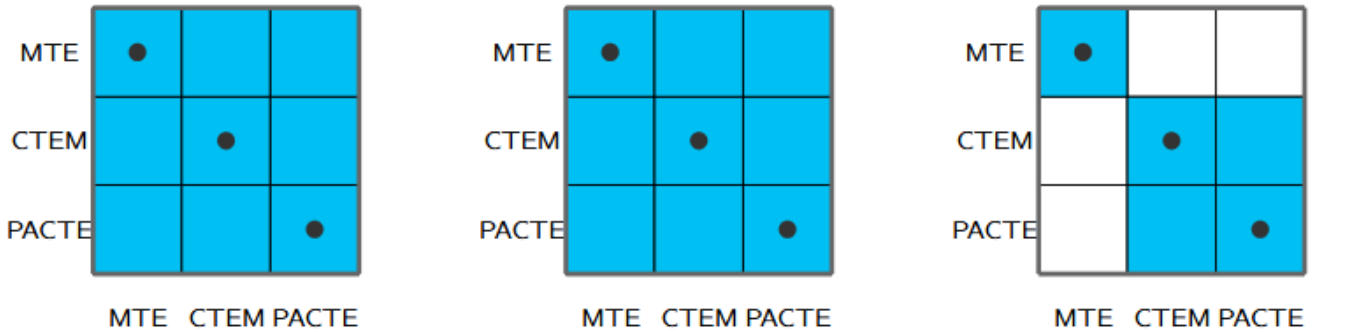
Assumptions: the CATE function is known and the scale of the estimand corresponds to that used for modeling, e.g., (log) odds ratio for binary outcomes

Homogeneous (constant) CATE

$$E(Y^t \mid X) = g^{-1}\left(\beta_0 + \beta_X X + \beta_T t\right)$$

Table 2. Model-based marginal estimands for the homogeneous illustrative models. For count outcomes and the log link, person-time is assumed constant, such that the log rate ratio is collapsible and can be interpreted a log risk ratio.

Outcome	Link function	Summary measure	Marginal estimand
Continuous	Identity	Mean difference	Does not depend on the distribution of purely prognostic covariates
Count	Logarithmic	Log risk ratio	Does not depend on the distribution of purely prognostic covariates
Binary	Logit	Log odds ratio	Depends on the full joint distribution of purely prognostic covariates



(a) Identity link, mean difference **(b)** Logarithmic link, log risk ratio **(c)** Logistic link, log odds ratio

Figure 1. Matrices indicating whether different estimands are equivalent for the homogeneous illustrative models. The blue squares denote matching estimand values; the dots denote the diagonal, where estimands are equivalent by definition.

Take-home message

For the (log) odds ratio, which is non-collapsible:

Even in the absence of effect modification…

- The MTE does not equal the CTEM or the PACTE
- The MTE depends on the full joint distribution of purely prognostic variables

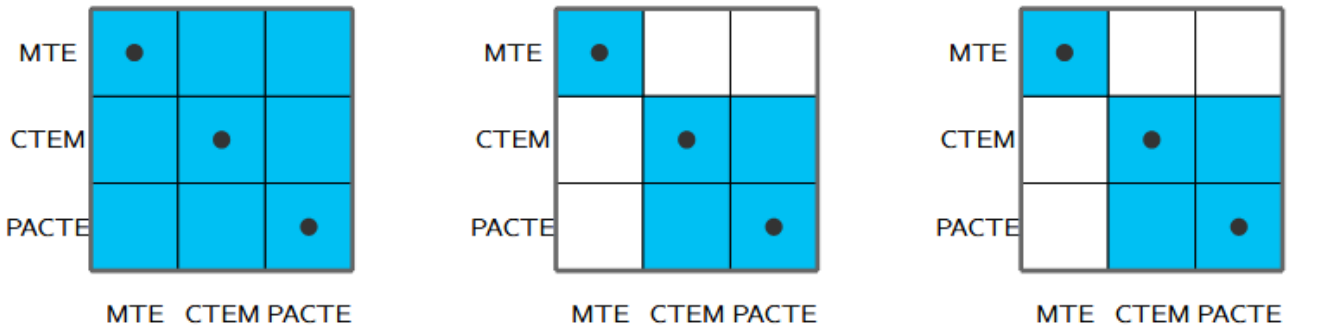
Remiro-Azócar, A., Phillippo, D.M., Welton, N.J., Dias, S., Ades, A.E., Heath, A. and Baio, G., 2025. Marginal and conditional summary measures: transportability and compatibility across studies. arXiv preprint arXiv:2507.21925.

Linear heterogeneous CATE

$$E(Y^t \mid X) = g^{-1}\left(\beta_0 + \beta_X X + \beta_T t + \beta_{XT} X t\right)$$

Table 3. Model-based marginal estimands for the heterogeneous illustrative models. For count outcomes and the log link, person-time is assumed constant, such that the log rate ratio is collapsible and can be interpreted as a log risk ratio.

Outcome	Link function	Summary measure	Marginal estimand
Continuous	Identity	Mean difference	Only depends on effect modifier means
Count	Logarithmic	Log risk ratio	Depends on the full joint distribution of effect modifiers and purely prognostic covariates that are associated with the former
Binary	Logit	Log odds ratio	Depends on the full joint distribution of effect modifiers and purely prognostic covariates



(a) Identity link, mean difference **(b)** Logarithmic link, log rate ratio **(c)** Logistic link, log odds ratio

Figure 2. Matrices indicating whether different estimands are equivalent for the heterogeneous illustrative models. The blue squares denote matching estimand values; the dots denote the diagonal, where estimands are equivalent by definition.

Take-home message

For the (log) risk ratio, which is collapsible but not directly collapsible:

Under the presence of effect modification…

- The MTE does not equal the CTEM or the PACTE
- The MTE depends on the full joint distribution of purely prognostic variables

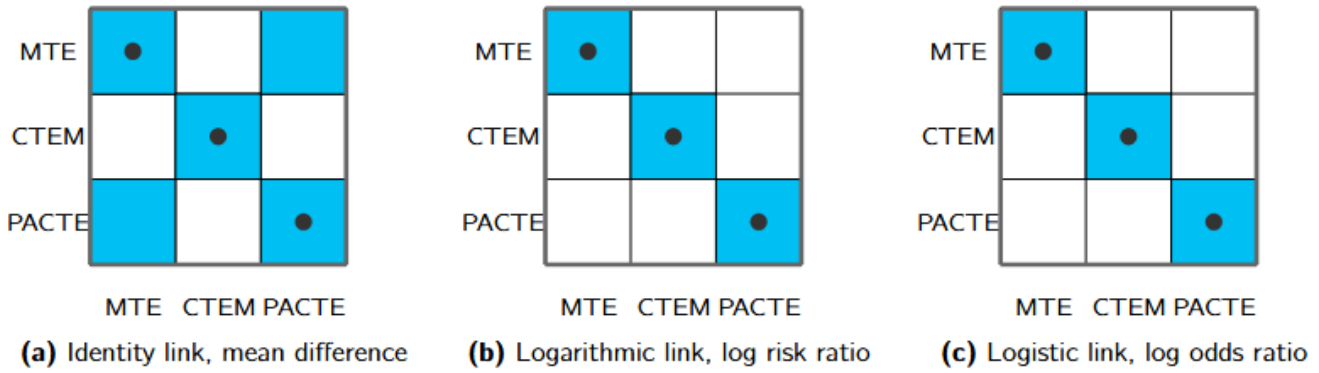
Remiro-Azócar, A., Phillippo, D.M., Welton, N.J., Dias, S., Ades, A.E., Heath, A. and Baio, G., 2025. Marginal and conditional summary measures: transportability and compatibility across studies. arXiv preprint arXiv:2507.21925.

Non-linear heterogeneous CATE

$$E(Y^t \mid X) = g^{-1}\left(\beta_0 + \beta_1 X + \beta_2 X^2 + \beta_T t + \beta_{1T} X t + \beta_{2T} X^2 t\right)$$

Table 4. Model-based marginal estimands for the quadratic (heterogeneous) illustrative models. For count outcomes and the log link, person-time is assumed constant, such that the log rate ratio is collapsible and can be interpreted as a log risk ratio.

Outcome	Link function	Summary measure	Marginal estimand
Continuous	Identity	Mean difference	Depends on effect modifier means and variances
Count	Logarithmic	Log risk ratio	Depends on the full joint distribution of effect modifiers and purely prognostic covariates that are associated with the former
Binary	Logit	Log odds ratio	Depends on the full joint distribution of effect modifiers and purely prognostic covariates



(a) Identity link, mean difference **(b)** Logarithmic link, log risk ratio **(c)** Logistic link, log odds ratio

Figure 3. Matrices indicating whether different estimands are equivalent for the quadratic (heterogeneous) illustrative models. The blue squares denote matching estimand values; the dots denote the diagonal, where estimands are equivalent by definition.

Take-home message

For all summary measures:

Under the presence of effect modification and where the CATE varies non-linearly with covariates…

- The PACTE does not equal the CTEM

Remiro-Azócar, A., Phillippo, D.M., Welton, N.J., Dias, S., Ades, A.E., Heath, A. and Baio, G., 2025. Marginal and conditional summary measures: transportability and compatibility across studies. arXiv preprint arXiv:2507.21925.

Transportability: marginal versus conditional

Marginal treatment effect:

- Depends on the distribution of “baseline risk” and observed prognostic factors
- Estimand (summary measure) does not depend on the covariate adjustment set
- Can be identified from an ideal RCT with minimal assumptions, even if adjusting for covariates

Conditional treatment effects:

- Do not depend on the distribution of baseline risk/purely prognostic factors
- Estimand (summary measure) changes with the covariate adjustment set
- Identification within an RCT may require statistical assumptions about model validity

While within-trial estimation of the marginal estimand may require weaker statistical assumptions, it may require stronger assumptions for transportability across studies/populations!

What about the hazard ratio?

Even for a constant CATE on the (log) hazard ratio scale…

…the marginal (log) hazard ratio depends on the shape of the hazard function, the distributions of the baseline hazard and observed purely prognostic factors, the length of follow-up and observed censoring pattern!

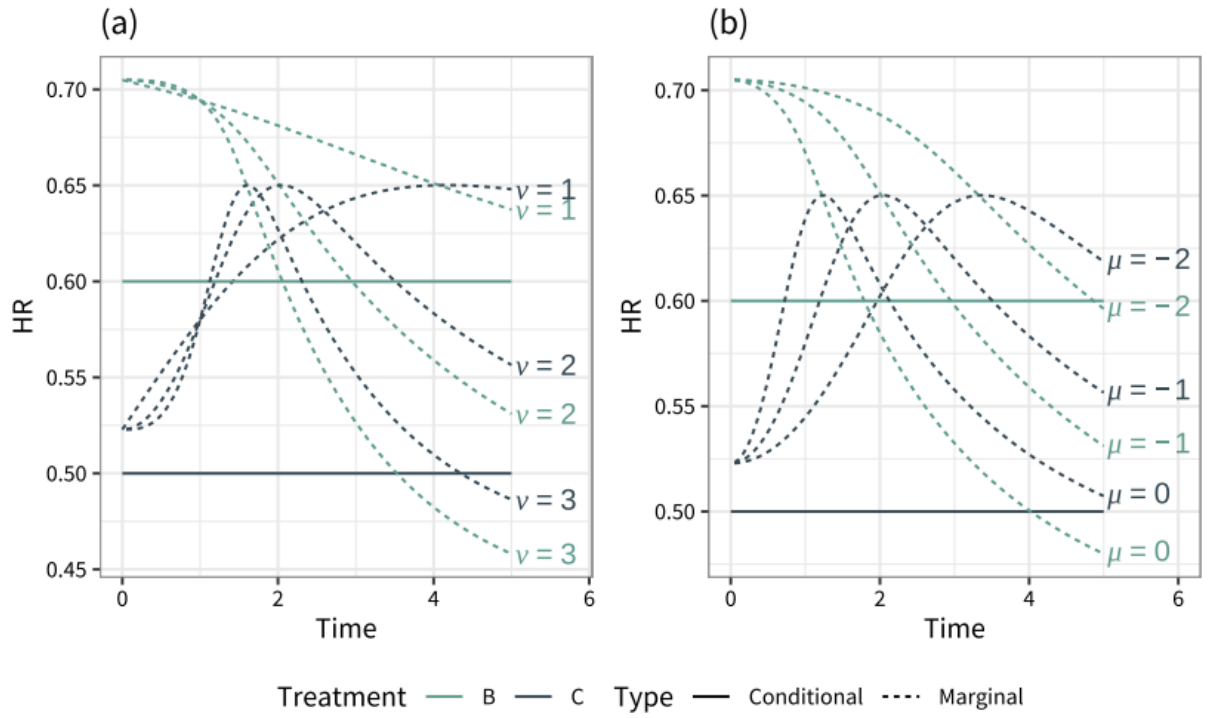


Figure 8. *Population-average conditional and marginal hazard ratios vs. treatment A over time, varying (a) the shape of the baseline hazard function $\nu_{(P)}$, and (b) the distribution of baseline log hazard $\mu_{(P)}$.*

Phillippo, D.M., et al., 2025. Effect modification and non-collapsibility together may lead to conflicting treatment decisions: A review of marginal and conditional estimands and recommendations for decision-making. Research Synthesis Methods, 16(2), pp.323-349.

Implications for evidence synthesis

Aggregate-level data meta-analysis

Meta-analysis of marginal effects:

- Validity is compromised if trial populations are heterogeneous
- Reasonable if the population in the scope is narrowly defined

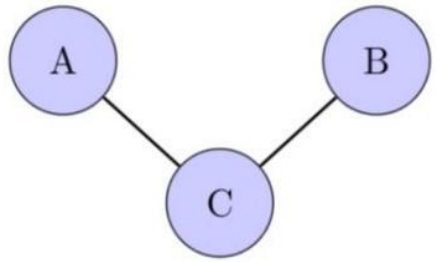
Meta-analysis of conditional effects:

- Validity is compromised if summary measures condition on different covariate sets
- …may also be compromised if populations are heterogeneous in observed or unobserved effect modifiers

Full access to IPD provides solutions:

- Pool adjusted marginal estimates that have been transported to the same target population
- Meta-regression (targets conditional effect) followed by standardization (targets marginal effect) – can produce any desired marginal or conditional summary measure in any target population

Population-adjusted indirect comparisons



Population-adjusted indirect comparisons are increasingly used to adjust for differences in effect modifiers between studies

Pairwise approaches: MAIC (weighting), STC/G-computation (outcome modeling)

$$\hat{\Delta}_{AB}^{(BC)} = \hat{\Delta}_{AC}^{(BC)} - \hat{\Delta}_{BC}^{(BC)}$$

Beware of estimand incompatibility issues in pairwise approaches, particularly for outcome modeling-based population adjustment!

Multilevel network meta-regression (ML-NMR)

Table 5. Summary measures estimated by different covariate adjustment approaches in the context of indirect comparisons.

Methodology	Summary measure
Matching-adjusted indirect comparison (MAIC)	MTE
"Plug-in" simulated treatment comparison (STC-P)	CTEM
"G-computation" simulated treatment comparison (STC-G)	MTE
Multilevel network meta-regression (ML-NMR)	PACTE and MTE
Network meta-interpolation (NMI)	CTEM

Remiro-Azócar, A., Phillippo, D.M., Welton, N.J., Dias, S., Ades, A.E., Heath, A. and Baio, G., 2025. Marginal and conditional summary measures: transportability and compatibility across studies. arXiv preprint arXiv:2507.21925.

ML-NMR:

- Not susceptible to estimand incompatibility issues (combines evidence at the CATE level)
- Can produce marginal and conditional population-average treatment effect estimates…
- …in any decision-relevant target population, (typically) under an additional identifying assumption

UИORDINARY
DRIVES
CHANGE

novo nordisk®

Thank you!