Research Synthesis Methods (2026), **00**: 1–20
doi:10.1017/rsm.2026.10098

Research Synthesis Methods

RESEARCH ARTICLE

Model-based network meta-analysis: Joint estimation of dose–response and time–course relationships

Anders Strathe[1], Martin Bøg[2] and Anders Gorst-Rasmussen[3]

[1]Pharmacometrics, Novo Nordisk A/S, Denmark
[2]HTA biostatistics, Novo Nordisk A/S, Denmark
[3]HTA Biostatistics, Novo Nordisk A/S, Denmark

Corresponding author: Anders Strathe; Email: aqss@novonordisk.com

Received: 29 April 2025; **Revised:** 13 March 2026; **Accepted:** 21 March 2026

Keywords: model-based; network meta-analysis; longitudinal; doseresponse; bayesian

Abstract

Under the network meta-analysis (NMA) framework for aggregate data, there are limited possibilities for evidence synthesis across multiple treatment doses or timepoints. Model-based network meta-analysis (MBNMA) has been recommended as a framework for either evidence synthesis across multiple dose levels or across multiple timepoints to circumvent the limitations of the NMA. A joint dose–response and time–course MBNMA (DT-MBNMA) is proposed that combines the strengths of both DT-MBNMA. This framework allows for combining data at multiple timepoints from studies in early clinical development with a broad range of doses to late-stage clinical studies with a limited range of doses. The method respects randomization and allows for assessment of consistency and hence satisfies the requirements from reimbursement agencies. The method was validated in a simulation study, showing that the drug effect parameters and therefore indirect treatment effects could be recovered without bias, while the precision of the treatment effects was dependent on the simulated network. Compared with a standard NMA, the methodology increased the statistical efficiency of the indirect treatment comparison (ITC). The use of the method was further illustrated on a dataset consisting of seven randomized clinical trials (RCTs) (26 treatment arms) in the treatment of obesity with Glucagon-like peptide-1 receptor agonists (GLP-1 RAs), with a broad range of doses and follow-up times. The method integrated phase 2 and 3 data seamlessly into the meta-analysis and provided greater precision on the treatment effects compared to NMA. Finally, the statistical framework may be used to support clinical decision-making, providing a framework for ITC during drug development.

Highlights

What is already known?

- Network meta-analysis (NMA) has limitations for synthesizing evidence across multiple treatment doses and timepoints, which can lead to disconnected treatment networks, when available data are discarded, or treatment nodes are pooled.
- Model-based network meta-analysis (MBNMA) has been developed to synthesize information across either multiple dose levels or multiple timepoints.

What is new?

- We extend the MBNMA framework to a joint dose–response and time–course MBNMA (DT-MBNMA), combining the strengths of DT-MBNMA.

This article was awarded Open Data and Open Materials badges for transparent practices. See the Data availability statement for details.

© The Author(s), 2026. Published by Cambridge University Press on behalf of The Society for Research Synthesis Methodology. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

- We use both simulations and a network of studies in obesity, to demonstrate how DT-MBNMA seamlessly integrates phase 2 and phase 3 data in NMA and provides greater precision on the indirect treatment effects, as compared to standard NMA.

Potential impact for *RSM* readers

- Phase 2 dose-finding studies can offer valuable information for health technology assessment (HTA), because the studies often include an active comparator, but their inclusion in standard evidence synthesis is often hampered by shorter follow-up and/or investigational doses.
- DT-MBNMA provides comprehensive integration of data generated across a drug development program, with applications to inform trial designs and HTA.

1. Introduction

Network meta-analysis (NMA) plays a crucial role in health technology assessment (HTA) by providing methods to compare multiple interventions simultaneously when direct head-to-head comparisons between all interventions are not available from individual randomized clinical trials (RCTs). NMA respects randomization and combines relative treatment effects, estimating heterogeneity between studies and is able to provide comparison between any two treatments in a connected network. However, the NMA framework has limitations in synthesizing evidence across multiple treatment doses and timepoints, which can lead to disconnected networks.[1]

Model-based meta-analysis (MBMA) is a recent method that addresses some of these limitations by handling complex dose–response and time–course relationships constructed from RCTs in the form of aggregate data.[2] The approach centres around modelling of arm-level results (i.e., absolute effects). In MBMA, aggregated data from individual studies are integrated into a single statistical model, which allows for the estimation of placebo, drug response, and assessment of heterogeneity across arm-level results. However, this approach models absolute outcomes in each trial arm rather than within-trial treatment contrasts, thereby foregoing the protection against confounding that randomization provides for estimation of treatment effects.

Recently, model-based network meta-analysis (MBNMA) has been proposed to combine both approaches,[3] respecting randomization while borrowing strength across multiple dose levels of treatments or timepoints. Pedder et al.[3] developed time–course MBNMA to model treatment effects over time, treating each treatment (or dose) as a discrete entity. Separately, Mawdsley et al.[4] developed dose–response MBNMA to model continuous dose–response relationships at single timepoints. Our framework extends these approaches by jointly modelling both dimensions, allowing structural parameters of time–course functions to themselves be functions of dose. The framework allows for combining data at multiple timepoints from studies in early clinical development with a broad range of doses to late-stage clinical studies with a limited range of doses. This may increase statistical efficiency by borrowing strengths from the full dose–response profile.[1] The dose–response and time–course MBNMA (DT-MBNMA) framework reduces to time–course MBNMA (T-MBNMA) when dose is modelled as discrete treatment effects and to dose–response MBNMA when analysing a single timepoint.

Borrowing strength across multiple studies in this way has applications in HTA, where careful integration of phase 2 dose-finding studies can provide valuable additional evidence. Phase 2 studies often include active comparators and broader dose ranges, but have shorter treatment durations than phase 3. The DT-MBNMA framework accounts for these differences in treatment duration and dose selection, enabling evidence synthesis across development phases. We follow the MBNMA literature and focus on the typical situation in HTA where only aggregate summary data are available. Beyond HTA, the framework's ability to predict both relative and absolute effects at any dose and timepoint also makes it a useful tool for decision-making and planning within pharmaceutical companies.

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

We illustrate the DT-MBNMA framework with an application in obesity. Obesity is a chronic disease with high prevalence and associated comorbidities, making it a growing global health concern. Comorbidities, such as type 2 diabetes, hypertension, ventilatory dysfunction, arthrosis, venous and lymphatic circulation diseases, and depression underscore the wide-ranging health implications of obesity, resulting in increased morbidity and mortality.[5] Here, the treatment duration required for the stabilization of weight loss profile is long and therefore early clinical development (e.g., dose-finding studies) does not explore the full treatment potential of a given dose.[6]

The DT-MBNMA framework is first evaluated in a simulation study, where the true values and model structure are known and then second applied to a network of RCTs in the therapeutic area of obesity disease. We end with a discussion of the framework in the context of guiding future study designs and supporting clinical decision-making.

2. Methods

2.1. Overview of modelling

The DT-MBNMA framework addresses how treatment effects evolve over time and vary across doses by combining three components:

1. Response model (Section 2.2): The model uses aggregate data extracted from publications (means and standard errors [SEs] at each timepoint for each study arm). These summaries are assumed to follow a multivariate normal distribution, with expected values come from parametric functions describing time–course and dose–response (Table 1). SEs are treated as known, extracted from publications while correlations are either set to zero, estimated using simple time series models, or uses observed correlations.
2. Parameter model (Section 2.3): The model estimates structural parameters that govern response functions rather than separate effects for each dose-timepoint combination. Response functions are defined based on subject-matter knowledge. For example, a time–course might be characterized by maximum effect and rate of onset; dose–response by maximum effect and the dose producing half-maximal response. These parameters vary across studies through study-specific reference levels (fixed effects) and treatment-specific deviations (random effects). Dose—response functions map dose levels to relative effects, enabling information sharing across doses, timepoints, and trials.
3. Random effects specification and diagnostics (Section 2.4): Between-study variability in treatment effects is handled through random effects with a covariance matrix that allows parameters to correlate (e.g., treatments differing in maximum effect may also differ in time-to-effect). Multi-arm trials receive special handling to account for correlation between arms sharing a common reference. This random effects structure provides diagnostics: large between-study variability suggests model misspecification or network inconsistency.

The framework nests simpler approaches: replacing parametric dose–response with discrete treatment effects yields T-MBNMA; using only end-of-treatment timepoints yields standard NMA. This nested conceptualization allows assessment of whether added complexity improves evidence synthesis.

2.2. Response model

We consider the typical setting where only aggregate-level trial data are available from publications or trial reports. Let $\mathbf{y}_{ij}$ denote a vector of summary statistics for study ($i = 1, \ldots, k$) and arm ($j = 1, \ldots, n_i$) across the time–course profile ($t = [0, \ldots, \tau_{ij}]$) with dimension ($n_i \times (\tau_{ij}+1)$). Here n_i denotes the number of arms in study i, and τ_{ij} denotes the end-of-treatment timepoint in study i and treatment arm j. Our framework applies to outcomes measured repeatedly at protocol-defined timepoints where time–course relationships are biologically plausible. Repeated observations on the individual are correlated over

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

Table 1. *Response functions for joint time–course and dose–response modelling.*

Model	Function: f(.)	Parameters	Time–course	Dose–response
Log-linear	$b \times \log(x + 1)$	Normal prior on b		x
Linear	$b \times x$	Normal prior on b		x
$E_{\max}$	$E_{max} \frac{x}{\exp(x_{50})+x}$	Normal priors on $E_{\max}$ and $\log(x_{50})$	x	x
Exponential	$E_{max}\,(1 - \exp(-\exp(kT) \times x))$	$E_{\max}$ and log(kT)	x	
Bateman	$A\left(\frac{\exp(kT_1)}{\exp(kT_1)-\exp(kT_2)}\right)\left(\exp\left(-\exp(kT_2) \times x - \exp\left(kT_1\right) \times x\right)\right.$	A, $\log(kT_2)$ and $\log(kT_1)$	x	

Note: The mark 'x' denotes function that explored for modelling time–course, dose–response or both.

time and therefore it is reasonable to specify a multivariate normal likelihood for the aggregate data:

$$\mathbf{y}_{ij} \sim N\left(\boldsymbol{\mu}_{ij}, \boldsymbol{\Sigma}_{ij}\right), \tag{1}$$

where $\boldsymbol{\mu}_{ij}$ is the vector of modelled outcomes generated from the response functions in Table 1, which are motivated in Section 2.2 of the article. The covariance matrix is denoted by $\sum_{ij}$. If the clinical outcome is given in absolute values, then an intercept should be added to the equations in Table 1, addressing the outcome at time = 0.

The diagonal elements in $\sum_{ij}$ are the within-arm variances ($\mathbf{se}^2_{ij}$), and in line with standard NMA, are assumed known. The within-arm SEs can be extracted or digitized from publications, whereas within-arm correlations seldom are reported. Therefore, we follow the approach of Pedder et al.[3] by either setting within-arm correlations to zero (i.e., a univariate normal likelihood) or estimating within-arm correlations based on the aggregate response summaries. Based on the work of Pedder et al.,[3] an autoregressive of first order structure (AR-1) appears reasonable because covariances are dependent on the time distance between clinical visits. In this article, data are simulated at the individual patient level (IPD). In the illustrative dataset, IPD is available to investigate the consequence of ignoring or approximating the within-arm correlation structure because correlations estimated from IPD will be at the individual level, which may be different to summary-level correlations.

2.3. Parameter model

In statistical terms, a random effects DT-MBNMA is a nonlinear mixed effect (NLME) model. NLME models are widely used in pharmacometrics, particularly for characterizing the time–course of drug concentrations and their effects on the body. The variation within and between studies is modelled through the structural parameters (e.g., time-constant (kT) or maximum effect ($E_{\max}$T) for the exponential function) of the response functions in Table 1, which are presented in detail in Section 2.4. To generalize, the model for the p'th structural parameter ($\theta_{ij}^{(p)}$) can be written as

$$g\left(\theta_{ij}^{(p)}\right) = \begin{cases} \mu_i^{(p)}, & j = 1 \ (\textit{reference}), \\ \mu_i^{(p)} + f\left(\mathbf{X_{ij}}\boldsymbol{\beta}^{(p)}, d_{ij}\right) + \gamma_{ij}^{(p)}, & j > 1, \end{cases} \tag{2}$$

where g(.) is a link function, e.g., the time-constants are log transformed, ensuring that they are positive real numbers on the time scale. The $\mu_i^{(p)}$ is the study-specific level for the reference treatment in the *i*th study, which is treated as a fixed effect and can be viewed as a nuisance parameter. The function f(.) is a parametric dose–response function and commonly used response functions are presented in Section 2.2 of the article. The response function f(.) maps the treatment doses to relative treatment effects and $\mathbf{X_{ij}}$ is a design matrix, encoding agent specific effects, and $\boldsymbol{\beta}^{(p)}$ is a vector of agent specific dose–response parameters (e.g., median effective dose (ED50) and maximum effect ($E_{\max}$D)). The proposed framework has the flexibility to model time–course profiles that are dose-dependent.

2.4. Random effects specification and diagnostics

The $\gamma_{ij}^{(p)}$ are the study-specific treatment effects, which are assumed to be exchangeable and handled as random effects. If $\gamma_{ij}^{(p)}$ are set to zero or dropped, then equation (2) reduces to a fixed effects model. Let $\boldsymbol{\gamma}_i$ be of vector of study-specific random effects of length n_{i}–1 x n_{p}, where n_{i} is the number of arms of study i as before, and n_{p} is the number of structural parameters (e.g., 3 for the Bateman equation, Table 1). Then, it is assumed:

$$\boldsymbol{\gamma}_i \sim N\left(\mathbf{0}, \boldsymbol{\Omega} \otimes \mathbf{T}_i\right), \tag{3}$$

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

where $\mathbf{\Omega}$ is an unstructured covariance matrix (n_p x n_p) for parameters, $\mathbf{T}_i$ is the known covariance matrix (n_i–1 by n_i–1) with 1 on the diagonal elements and 1/2 on the off-diagonal elements, accounting for correlated random effects in multi-arm studies.[7] The Kronecker product is denoted by ⊗.

The framework outlined in equation (2) accommodates T-MBNMA,[3] where the agent-by-dose combinations are viewed as separate treatments. If the parametric dose–response function f(.) in equation (2) is replaced by a vector of relative treatment effects for the agent-by-dose combinations, then the model corresponds to a T-MBNMA.[3] This relaxes the assumptions around the parametric form of the dose–response, i.e., monotonicity, but does assume consistency. The random effects specification provides diagnostics for assessing dose–response model fit and the consistency assumption. Misspecification of the dose–response relationship increases between-study standard deviations (SDs) for the structural parameters defining treatment effects, providing a useful diagnostic for selection of an appropriate dose–response function.[8] Between-study heterogeneity, quantified through Ω, reflects variability in structural parameters not explained by the specified dose–response and time–course models.

Consistency can be assessed by comparing model fit and between-study SDs across nested model specifications. Fitting all direct connections in the network extends the framework to the unrelated mean effects model (UME), relaxing both dose–response assumptions and the consistency assumption.[8] Substantially better model fit or lower between-study SDs in the UME is suggestive of inconsistency. Deviance–deviance plots comparing the parametric models to the UME provide a graphical diagnostic for detecting inconsistency. Stable between-study SDs when moving from UME to T-MBNMA to DT-MBNMA, with points clustering near the line of equality in deviance–deviance plots, indicates that parametric restrictions are compatible with the data and the consistency assumption is plausibly satisfied.

Residual heterogeneity following dose–response and time–course modelling may indicate effect modification by unmeasured covariates, suggesting potential transitivity violations. Similar to NMA, prior to analysis, the exchangeability of study populations should be assessed in order to reduce potential transitivity violations. Visual predictive checks, a common method used to diagnose and evaluate pharmacometric models,[9] provides a complementary assessment of whether the model captures observed data patterns across all study arms.

2.5. Response function specification and model selection

Parametric response functions are used to link the relationship between clinical outcome to time and dose. The response functions used in the paper are listed in Table 1. Throughout the article 'T' is used to denote a time–course parameter and 'D' denotes a dose–response parameter. The listed functions (Table 1) include the E_{max} function, which is widely used in pharmacology and drug development to understand and quantify the relationship between drug dose/concentration and effect, aiding in the determination of optimal dosages and the prediction of drug responses under different doses/concentrations. It has two parameters E_{max}D and ED50, which are the maximum effect (as dose $\rightarrow\infty$) and the median effective dose. The log-linear and linear functions are obvious choices if efficacy endpoints are far from plateauing for incremental doses. The framework is not limited to the response functions considered here.

Time–course relationships are usually characterized by a nadir in treatment response and the time delay between treatment initiation and nadir. This is conveniently expressed using functional forms that are parametrized with these quantities. The E_{max} function has been applied,[3] but also the exponential function captures the relationship, where the time–course parameters are denoted E_{max}T (as time $\rightarrow\infty$) and ET50 (time to 50% maximum response). The time-constant in the exponential function (i.e., kT) can be expressed in terms of a pharmacological half-life (i.e., log(2)/kT) and the effect, when time $\rightarrow\infty$ as E_{max}T, representing the nadir in treatment response. Finally, we introduce the Bateman equation, which is a function of two exponentials with a biphasic pattern and is well known from pharmacokinetics[10] (i.e., one compartment model with linear absorption and elimination). The parameters can be associated with observable clinical events such as improvement in response and return to baseline phenomena

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

(e.g., due to disease progression or trial fatigue). The time-constants can be expressed in terms of half-lives (i.e., $\log(2)/kT_1$ and $\log(2)/kT_2$) like with the exponential function. For model comparison, the expected log posterior density (ELPD) criterion is used as defined in the loo package in R,[11] where larger criterion values reflect a better-performing model. It approximates the leave-one-out cross-validation and has been shown to have various advantages over simpler estimates of predictive error such as the deviance information criterion (DIC).[11] Moreover, an approximate SE for estimated predictive errors can also be obtained, which is useful for comparison of predictive errors between two or more competing models.

2.6. Simulation study

The following section describes a simulation study undertaken to understand the robustness of MBNMA under different scenarios and to compare it to standard NMA and T-MBNMA in terms of the gain in precision of relative treatment effects. We emulate the situation where phase 2 data are included in the evidence synthesis.

A network of three agents (q = [A, B, C]) is simulated. We assume a typical development program, where one phase 2 dose-finding trial is conducted, having a placebo/standard-of-care arm and four arms with increasing dose levels. This is followed by two phase 3 confirmatory trials with a placebo/standard-of-care arm and two treatment arms. Finally, a phase 3 trial is conducted that provides direct evidence of A. against B. A table with design features is given in Table S1 in the Supplementary Material, and the corresponding network is drawn in Figure S1 in the Supplementary Material. The clinical outcome is modelled from an exponential function (i.e., $E_{max}T\,(1-\exp(-\exp(kT)\times t))$, Table 1), having two parameters $E_{\max}$T and kT. Further, it is assumed that these parameters change in dose-dependent manner, where $E_{\max}$T will change according to an $E_{\max}$ function and kT will change according to a log-linear function. Hence, the simulation model for the study-by-arm specific parameters can be written as

$$log\left(kT_{ij}\right) = \begin{cases} \mu_i, & j = 1\ (reference), \\ \mu_i + b^{(q)}\log\left(d_{ij}+1\right) + \gamma^1_{ij}, & j > 1, \end{cases} \tag{4}$$

$$E_{max}T_{ij} = \begin{cases} \mu_i, & j = 1\ (reference), \\ \mu_i + \frac{E_{max}D^{(q)}\times d_{ij}}{\exp\left(ED50^{(q)}\right)+d_{ij}} + \gamma^2_{ij}, & j > 1. \end{cases} \tag{5}$$

The study-specific treatment effects are assumed random and are generated from equation (3). The numerical values used in the simulations are given in Table S2 in the Supplementary Material. In the simulations, the long-term treatment effect is given as $E_{\max}$T in the exponential function (i.e., effect as time $\to\infty$). Therefore, indirect treatment comparisons (ITCs) are inferred from $E_{\max}$T in the T-MBNMA and DT-MBNMA, whereas NMA uses the week 52 study results, as NMA cannot extrapolate. The true mean time–course and dose–response profiles are presented in Figure S2A and S2B in the Supplementary Material. In Figure S2A in the Supplementary Material, the true mean time–course profile for higher doses of agent C, the time–course has not reached its nadir (i.e., stabilized) and hence by design the full effect size cannot be quantified unless the time–course treatment profile is used and modelled. This design feature is introduced to showcase the limitations of standard NMA (i.e., the week 52 timepoint) for handling a clinical endpoint that is evolving over time.

Three scenarios are studied, where treatment duration and study selection are varied. These are:

1. A case, where all studies and treatment arms are observed for 52 weeks. This is considered the base case.
2. A case, where phase 2 for agent C is truncated to 28 weeks. This is considered the realistic case, where the phase 2 study for agent C is conducted with a shorter treatment duration.
3. A case, where phase 2 is truncated to 28 weeks and phase 3 for agent C is excluded. This is a scenario that mimics an ITC feasibility assessment, where the proposed statistical framework is used to predict the long-term efficacy of agent C against agents A and B. Hence, the statistical

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

framework is used as part of dose setting for phase 3 development for agent C, i.e., model-informed drug development.

Two ITCs are assumed of interest, comparing 'agent A – dose 1 mg' vs. 'agent C – dose 2 mg' and 'agent B – dose 3 mg' vs. 'agent C – dose 2 mg'. For simplicity and ease of notation, ITCs are denoted 'A vs. C' and 'B vs. C', going forward.

The IPD are simulated at timepoints 2, 4, 8, 12, 16, 20, 28, 36, 44 and 52 weeks after randomization from a multivariate normal distribution with the mean given by the exponential response function and arm-level parameters specified by equations (4) and (5). The individual patient covariance matrix is presented in Table S3 in the Supplementary Material. Similar to published results, each treatment arm IPD within a clinical study is aggregated into a mean time–course profile and corresponding SE. In addition, the within-arm correlation matrix is also derived and used in the parameter estimation. Three analysis models are considered:

1. Ignore the within-arm correlation structure by specifying a univariate normal likelihood. Here, the within-arm correlation structure is set to zero.
2. Account for within-arm correlation structure by specifying a multivariate normal likelihood. Here, the within-arm correlation structure is approximated with an AR-1 process.
3. Account for within-arm correlation structure by specifying a multivariate normal likelihood. Here, the observed within-arm correlation matrix is used in the estimation.

We estimate parameters using the random effects DT-MBNMA with an exponential time–course function and the dose–response specified in equations (4) and (5). As a comparator estimation method, we also apply standard NMA, where absolute aggregated response at week 52 is used as the dependent variable. A T-MBNMA used as an additional comparator model, where absolute aggregated response at all timepoints is used as the dependent variable. For both analysis models, phase 2 is excluded. Simulation results are summarized in standard measures: Median absolute deviation (MAD), root mean square error (RMSE), and bias.

2.7. Illustrative example: A network of studies in obesity

Glucagon-like peptide-1 receptor agonists (GLP-1 RAs), used to treat type 2 diabetes, have been shown to be effective in generating weight loss in preclinical and clinical studies. Thus, the methodology is illustrated using a dataset of RCTs investigating obesity medication or obesity therapy with GLP-1 RAs. Studies used in the current analysis are published in peer-reviewed journals, comparing the effect of a GLP-1 RA to another GLP-1 RA or placebo (both as adjunct to lifestyle interventions) for weight reduction in patients with obesity. Although individual patient data were available, we analysed only aggregate summary statistics (means and SEs at protocol-defined timepoints) to illustrate the methodology as it would be applied to published trial reports where IPD is unavailable. GLP-1 RAs approved for weight management include Saxenda®[12] (active agent: liraglutide) and Wegovy®[13] (active agent: semaglutide). For Saxenda, the approved dosage is liraglutide 3.0 mg once-daily®[12] and for Wegovy it is semaglutide 2.4 mg and 1.7 mg once-weekly®.[13] The primary endpoint in weight management studies is the percent change in body weight (BW) from baseline to end of treatment, which is collected at selected timepoints during treatment. Neither the end-of-treatment timepoint nor the selected timepoints during treatment are aligned across studies (Table S4 in the Supplementary Material). A phase 2 dose-finding study[6] for liraglutide was conducted, investigating the dose–response of incremental liraglutide doses on weight and a large phase 3 trial[14] establishing safety and efficacy of liraglutide 3.0 mg once-daily. For semaglutide, a phase 2 dose-finding study[15] investigating the dose–response of increasing once-daily doses of semaglutide on weight, is available. Semaglutide is a GLP-1 analogue that has an elimination half-life of approximately 1 week,[16] supporting once-weekly dosing, and therefore a once-weekly dose regimen was selected for the phase 3 development program. Hence, the semaglutide once-weekly dose equivalents are computed by multiplying the once-daily doses in phase 2 with a factor of

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

7, which is supported by pharmacokinetic data collected in the trials.[17] Four phase 3 studies[18–21] are included, investigating safety and efficacy of 2.4 mg semaglutide and additionally one of the studies investigated safety and efficacy of 1.7 mg once-weekly.[20]

The treatment network is shown in Figure 1, where Figure 1a is the full treatment network and Figure 1a is reduced to phase 3 studies.

The modelling workflow is defined as follows:

1. Time–course models (Table 1) not including a dose–response component are fitted and ranked according to the goodness of fit using the ELPD criterion. First, common-effects models are fitted, followed by the introduction of exchangeable treatment effects, and finally an AR-1 process is introduced for the arm-correlation structure. The dose–response is specified as a set of discrete treatment effects (i.e., a treatment-level MBNMA), making no assumption about the shape of the dose–response, but does assume consistency.
2. Based on the best performing time–course model, parametric dose–response models are introduced into the structural parameter of the time–course functions. These models are compared to the treatment-level MBNMA and the UME, using graphical methods[8] and information criteria. For the selected dose–response model in this step, the observed within-arm correlations are introduced and used in the estimation of dose–response parameters.

Two standard NMAs are conducted for comparison purposes. The first uses primary statistical analysis (including imputation of missing data) that was defined in the individual study protocols and evaluated at the per protocol specified end-of-treatment timepoint (spanning from week 56 to week 104). The second uses the aggregated data (as observed) at the per protocol specified end-of-treatment timepoint, which is introduced to align with aggregated time–course data used in the DT-MBNMAs.

2.8. Implementation

Estimation and inference is carried out using Stan (Stan Development Team, version 2.26.1), where the R package brms provides an interface to R (version 2.21.1).[22] All computations are conducted with R version 4.3.1 (a GitHub repo is available at https://github.com/agorstras/dt-mbnma). Models that included the observed within-arm covariance matrix are implemented directly in Stan because these cannot be fitted with the current version of brms. Choice of prior distributions and convergence diagnostic are presented below and in-detail in the Supplementary Material.

2.9. Convergence monitoring and prior specification

In stochastic simulation and estimation part, two independent chains were run for 1,000 burn-in iterations followed by 1,000 iterations. Gelman's R statistic was monitored for the model parameters during the simulations, and it appeared that models converged as all parameters had an $R < 1.05$. In the analysis of the illustrative dataset, two independent chains were run for 2,500 burn-in iterations followed by 2,500 iterations. Gelman's R statistic was monitored for the model parameters during the simulations, and it appeared that models converged as all parameters had an $R < 1.05$. If the models did not converge, the number of iterations were tripled.

Vague normal prior distributions with a SD of 100 were given to the nuisance and drug effect parameters (e.g., b or x50, Table 1). For parameters that only take positive values (e.g., time-constants, median effective dose), priors were expressed on a log-scale. The between-study SDs on exchangeable parameters were given half-normal prior distributions with an SD of 5. The prior for correlation matrix for the exchangeable parameters uses the Lewandowski-Kurowicka-Joe prior (LKJ) is used with a degree freedom of 1. In models with a multivariate likelihood, the correlation parameter was given a uniform prior distribution (uniform(–1,1)) to reflect that outcomes at different timepoints within the same study arm are correlated.

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

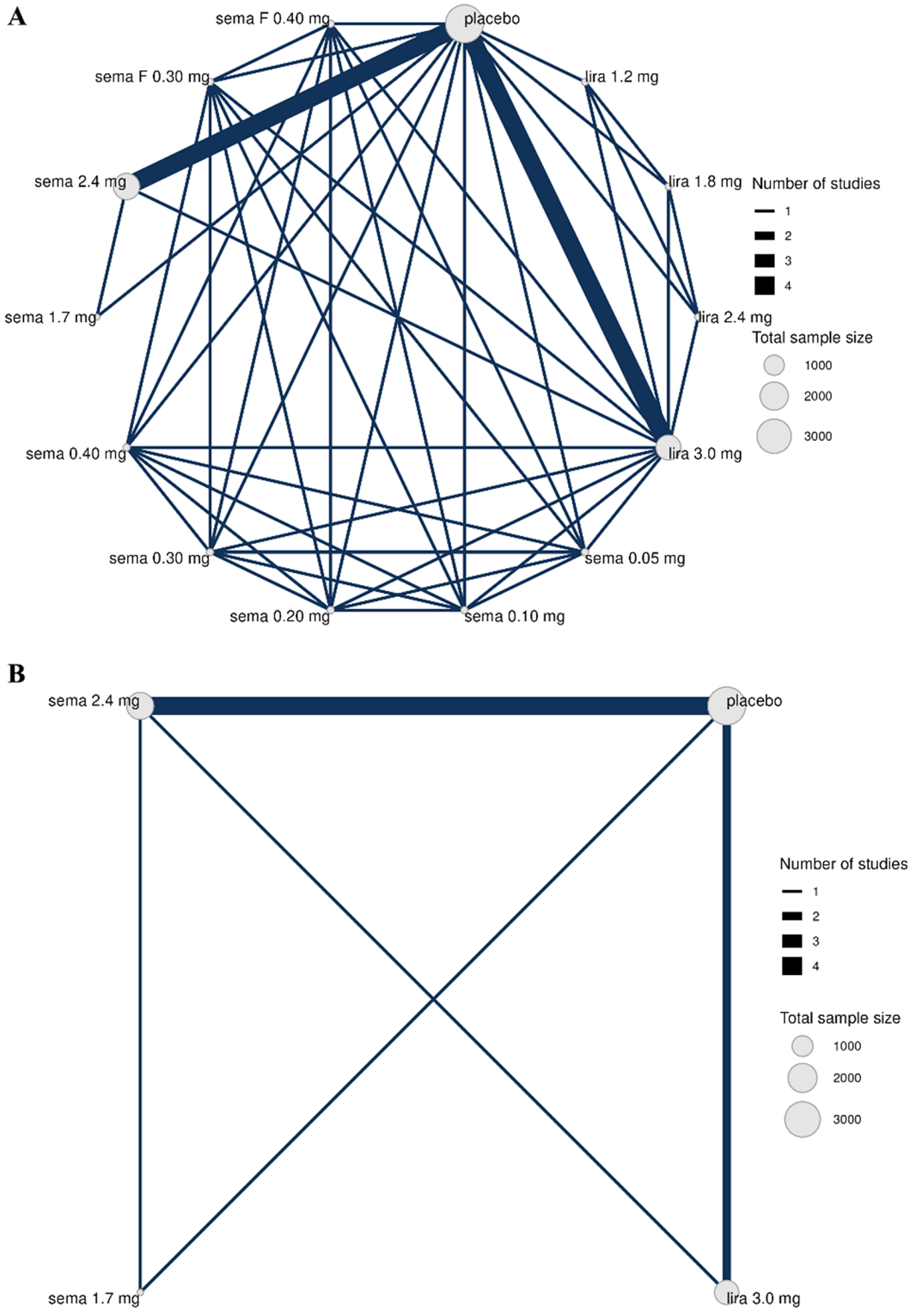


Figure 1. *Network of treatment comparisons from seven RCTs, where each treatment is represented by a node. Where direct RCT evidence exists for a particular comparison, the nodes are connected by a line, the thickness of which is proportional to the number of comparisons. Network 'a' is the full network, including phase 2 and 'b' includes only phase 3 data.*

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

Table 2. *Estimation of indirect treatment comparisons (ITC) under different scenarios: (1) All arms are observed for 52 weeks, (2) Truncate phase 2 for agent C to 28 weeks, and (3) Truncate phase 2 to 28 weeks and exclude phase 3 for agent C.*

Scenario	ITC	Model[a]	Mean	True	Bias	RMSE	MAD
(1)	A vs. C	Standard NMA	6.3	7.4	−1.0	1.2	1.1
(1)	A vs. C	Time–course dose–response	7.4	7.4	0.031	0.66	0.54
(1)	A vs. C	Time–course	7.4	7.4	0.018	0.65	0.53
(1)	B vs. C	Standard NMA	1.5	2.2	−0.72	1.0	0.88
(1)	B vs. C	Time–course dose–response	2.3	2.2	0.080	0.70	0.58
(1)	B vs. C	Time–course	2.2	2.2	0.013	0.75	0.60
(2)	A vs. C	Standard NMA	6.3	7.4	−1.0	1.2	1.1
(2)	A vs. C	Time–course dose–response	7.5	7.4	0.093	0.85	0.67
(2)	A vs. C	Time–course	7.4	7.4	0.018	0.65	0.53
(2)	B vs. C	Standard NMA	1.5	2.2	−0.72	1.0	0.88
(2)	B vs. C	Time–course dose–response	2.4	2.2	0.15	0.90	0.69
(2)	B vs. C	Time–course	2.2	2.2	0.013	0.75	0.60
(3)	A vs. C	Time–course dose–response	7.6	7.4	0.19	1.7	1.3
(3)	B vs. C	Time–course dose–response	2.5	2.2	0.24	1.7	1.3

Note: Independence within-arm across timepoints. RMSE, root mean square error; MAD, median absolute deviation.
[a]Individual patient data (arm-level) was simulated from an unstructured covariance matrix, generating correlated observations across timepoints. The mean time–course was simulated from an exponential function ($E_{max}T\ (1 - \exp(-\exp(kT) \times time))$), where the long-term treatment effect ($E_{\max}$T) is dose-dependent and described by an $E_{\max}$ model, having two parameters; maximum response relative to placebo ($E_{\max}D^{(q)}$) and the dose at which of half the maximum response can be achieved (ED50$^{(c)}$) for the 3 agents (q = [A, B, C]). Decay parameter (kT) was dose-dependent and described by a log-linear model, having 1 slope parameter (b$^{(c)}$). Estimation models (DT-MBNMA and T-MBNMA) assumes a univariate normal likelihood, where within-arm correlations are set to zero.

3. Results

3.1. Simulation study

Results of estimation of indirect treatment effects are shown in Tables 2–4. Table 2 depicts the case where within-arm correlations across timepoints are ignored, while Table 3 depicts the case where within-arm correlations across timepoints are approximated with an AR-1 process. Finally, Table 4 displays the case where observed correlations across timepoints are used in the estimation. The T-MBNMA and DT-MBNMA perform similar in terms of both bias and precision of the estimated indirect treatment effects. This is regardless of the correlation structure used in the estimation and across scenarios 1 and 2. Results of estimation of drug effect parameters from simulated data are shown in Figures S3–S5 in the Supplementary Material. Across scenarios 1 and 2, the drug effect parameters are on average estimated without bias because the medians of the individual point estimates (i.e., posterior means) are in line with the true values.

The standard NMA model shows lower precision of the estimated indirect treatment effects, but also underestimates the effect size of the indirect treatment effects, when the comparisons are based solely on week 52 data. The systematic underestimation of the effect size, comparing agents A and B to C is expected given the simulation design. In Figure S2A in the Supplementary Material, the true mean time–course profile for higher doses of agent C, the time–course has not reached its nadir (i.e., stabilized), and therefore the full effect size cannot be quantified unless the time–course profile is used and modelled.

In scenario 3, where the phase 2 treatment duration is truncated to 28 weeks and phase 3 data are not available, it appears that there is a minor underestimation of $E_{\max}D^{(c)}$ counteracted by a minor overestimation of drug potency ED50$^{(c)}$ for agent C. These findings are consistent regardless of the correlation structure used to model the underlying IPD. The impact on estimation of indirect treatment

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

Table 3. *Estimation of indirect treatment comparisons (ITC) under different scenarios: (1) All arms are observed for 52 weeks, (2) Truncate phase 2 for agent C to 28 weeks, and (3) Truncate phase 2 to 28 weeks and exclude phase 3 for agent C.*

Scenario	ITC	Model[a]	Mean	True	Bias	RMSE	MAD
(1)	A vs. C	Standard NMA	6.3	7.4	−1.0	1.2	1.1
(1)	A vs. C	Time–course dose–response	7.4	7.4	0.029	0.66	0.54
(1)	A vs. C	Time–course NMA	7.4	7.4	0.017	0.65	0.53
(1)	B vs. C	Standard NMA	1.5	2.2	−0.72	1.0	0.88
(1)	B vs. C	Time–course dose–response	2.3	2.2	0.080	0.70	0.58
(1)	B vs. C	Time–course NMA	2.2	2.2	0.018	0.74	0.60
(2)	A vs. C	Standard NMA	6.3	7.4	−1.0	1.2	1.1
(2)	A vs. C	Time–course dose–response	7.4	7.4	0.011	0.61	0.50
(2)	A vs. C	Time–course NMA	7.4	7.4	0.017	0.65	0.53
(2)	B vs. C	Standard NMA	1.5	2.2	−0.72	1.0	0.88
(2)	B vs. C	Time–course dose–response	2.3	2.2	0.063	0.66	0.55
(2)	B vs. C	Time–course NMA	2.2	2.2	0.018	0.74	0.60
(3)	A vs. C	Time–course dose–response	7.6	7.4	0.18	1.7	1.3
(3)	B vs. C	Time–course dose–response	2.4	2.2	0.23	1.7	1.3

Note: AR-1 within-arm correlation across timepoints. RMSE, root mean square error; MAD, median absolute deviation.
[a]Individual patient data (arm-level) was simulated from an unstructured covariance matrix, generating correlated observations across timepoints. The mean time–course was simulated from an exponential function ($E_{max}T\ (1 - \exp(-\exp(kT) \times time))$), where the long-term treatment effect ($E_{\max}$T) is dose-dependent and described by an $E_{\max}$ model, having two parameters; maximum response relative to placebo ($E_{\max}D^{(q)}$) and the dose at which of half the maximum response can be achieved ($ED50^{(c)}$) for the 3 agents (q = [A, B, C]). Decay parameter (kT) was dose-dependent and described by a log-linear model, having 1 slope parameter ($b^{(c)}$). Estimation models (DT-MBNMA and T-MBNMA) assumes multivariate normal likelihood, where within-arm correlations are estimated from aggregated data, using an AR-1 process.

effects is minimal as the effect sizes agree with the true values used in the simulations. As expected, the precision on the indirect treatment effects (as measured by RMSE or MAD) is halved, having only phase 2 data available for the ITC. Moreover, it reflects both a smaller study population in phase 2 and a shorter treatment duration.

3.2. Application to the obesity treatment network

Visual inspection of the data shows that the percent change in BW decreases in a nonlinear time-dependent fashion (Figure 2). In some of the placebo-treated groups, a trend towards baseline levels is observed. This suggests that a simple linear model will not be a good fit for the data, but that $E_{\max}$, exponential, or the Bateman models may be more suitable. Over time, a clear separation of the active treatment arms, showing that the relative effect size of semaglutide and liraglutide increases in a dose-dependent manner (Figure 2).

Table 5 gives an overview of the performance of the explored models. The T-MBNMA shows a minor differentiation between $E_{\max}$ and exponential models. The Bateman model had convergence problems when exchangeable treatment effects are introduced and therefore not reported. Introducing random effects into the time–course models, facilitating estimation of between-study SD in the structural parameters, shows improvements in the overall fit. Adding within-arm correlations structure in the form an AR-1 process significantly improved the fit to the data with the exponential model performing best. Hence, dose–response models are explored within the exponential time–course model, assuming exchangeable treatment effects and within-arm correlations structure in the form of an AR-1 process. Estimates of the autocorrelation parameter are given in Table 5.

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

Table 4. *Estimation of indirect treatment comparisons (ITC) under different scenarios: (1) All arms are observed for 52 weeks, (2) Truncate phase 2 for agent C to 28 weeks, and (3) Truncate phase 2 to 28 weeks and exclude phase 3 for agent C.*

Scenario	ITC	Model[a]	Mean	True	Bias	RMSE	MAD
(1)	A vs. C	Standard NMA	6.3	7.4	−1.0	1.2	1.1
(1)	A vs. C	Time–course dose–response	7.4	7.4	0.011	0.61	0.50
(1)	A vs. C	Time–course	7.4	7.4	0.012	0.64	0.52
(1)	B vs. C	Standard NMA	1.5	2.2	−0.72	1.0	0.88
(1)	B vs. C	Time–course dose–response	2.3	2.2	0.042	0.63	0.51
(1)	B vs. C	Time–course	2.2	2.2	0.010	0.73	0.61
(2)	A vs. C	Standard NMA	6.3	7.4	−1.0	1.2	1.1
(2)	A vs. C	Time–course dose–response	7.5	7.4	0.10	0.63	0.51
(2)	A vs. C	Time–course	7.4	7.4	0.012	0.64	0.52
(2)	B vs. C	Standard NMA	1.5	2.2	−0.72	1.0	0.88
(2)	B vs. C	Time–course dose–response	2.3	2.2	0.12	0.68	0.54
(2)	B vs. C	Time–course	2.2	2.2	0.010	0.73	0.61
(3)	A vs. C	Time–course dose–response	7.4	7.4	0.038	1.5	1.2
(3)	B vs. C	Time–course dose–response	2.3	2.2	0.067	1.4	1.1

Note: Observed within-arm correlations across timepoints. RMSE, root mean square error; MAD, median absolute deviation.
[a]Individual patient data (arm-level) was simulated from an unstructured covariance matrix, generating correlated observations across timepoints. The mean time–course was simulated from an exponential function ($E_{max}T\ (1 - \exp(-\exp(kT) \times time))$), where the long-term treatment effect ($E_{\max}$T) is dose-dependent and described by an $E_{\max}$ model, having two parameters; maximum response relative to placebo ($E_{\max}D^{(q)}$) and the dose at which of half the maximum response can be achieved ($ED50^{(c)}$) for the 3 agents (q = [A, B, C]). Decay parameter (kT) was dose-dependent and described by a log-linear model, having one slope parameter ($b^{(c)}$). Estimation models (DT-MBNMA and T-MBNMA) assume multivariate normal likelihood, where observed within-arm correlations are used in the estimation.

Table 5. *Model fit statistics for time–course models, fitted to the obesity dataset.*

Function	Treatment effect	Within-arm correlation	ELPD	SE	Δ_{ELPD}	SE(Δ_{ELPD})
Exponential	Exchangeable	AR–1	−97.4	11.2	0	0
$E_{\max}$	Exchangeable	AR–1	−116	13.1	−19.1	4.39
$E_{\max}$	Exchangeable	Independent	−202	31.4	−105	26.5
$E_{\max}$	Common	Independent	−216	31.4	−119	26.8
Exponential	Exchangeable	Independent	−257	42.2	−160	40.7
Exponential	Common	Independent	−273	43.2	−176	41.9
Bateman	Common	Independent	−280	43.3	−182	42.1

Note: ELPD, expected log posterior density estimate; SE, standard error.

Embedding dose–response models for the relative treatment effects on the structural parameters of the exponential time–course model ($E_{\max}$T and kT) was successful (Table 6). The comparison of DT-MBNMAs to T-MBNMA shows that the dose–response can be captured in a continuous monotonic form with relatively few drug-effect parameters to estimate. The change in the ELPD (Δ_{ELPD}) and its corresponding SE indicates that the models are performing equally well.

The DT-MBNMAs show a minor differentiation between log-linear and $E_{\max}$ model in terms of the between-study SDs for $E_{\max}$T and Kt, and that the SDs were in line with the T-MBNMA and UME models. This indicates that there are no major violations of assumptions around the dose–response relationship and the consistency assumption. Fitting a linear dose–response model shows that the between-study SD on the time–course parameter ($E_{\max}$T) increases substantially, indicating that

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

Table 6. *Model fit statistics for dose–response time–course models, fitted to the obesity dataset.*

E_{max} T[a]	kT[a]	ELPD	SE	Δ_{ELPD}	SE(Δ_{ELPD})	Between-study SD (E_{max} T)	Between-study SD (kT)	Correlation (E_{max} T, kT)	Within-arm correlation
Linear	Linear	−97	11	0	0	1.2 [0.49; 2.1]	0.18 [0.039; 0.34]	0.026 [−0.90; 0.77]	0.86 [0.71; 0.95]
Log-linear	Log-linear	−97	11	−0.56	1.5	0.73 [0.073; 1.6]	0.18 [0.065; 0.33]	−0.011 [−0.92; 0.87]	0.84 [0.68; 0.95]
Treatment level	Treatment level	−97	11	−0.59	1.4	0.73 [0.038; 2.1]	0.17 [0.0052; 0.58]	0.049 [−0.95; 0.96]	0.86 [0.71; 0.96]
UME	UME	−97	11	−0.62	1.4	0.74 [0.041; 2.1]	0.16 [0.0053; 0.56]	0.069 [−0.94; 0.96]	0.86 [0.71; 0.95]
E_{max}	Log-linear	−98	11	−0.72	1.3	0.74 [0.075; 1.6]	0.18 [0.057; 0.32]	−0.15 [−0.94; 0.79]	0.85 [0.68; 0.95]

Note: ELPD, expected log posterior density estimate; SE, standard error; UME, unrelated mean effects model.

[a]The exponential function $E_{max}T\ (1 - \exp(-\exp(kT) \times time))$ is specified for the time–course relationship.

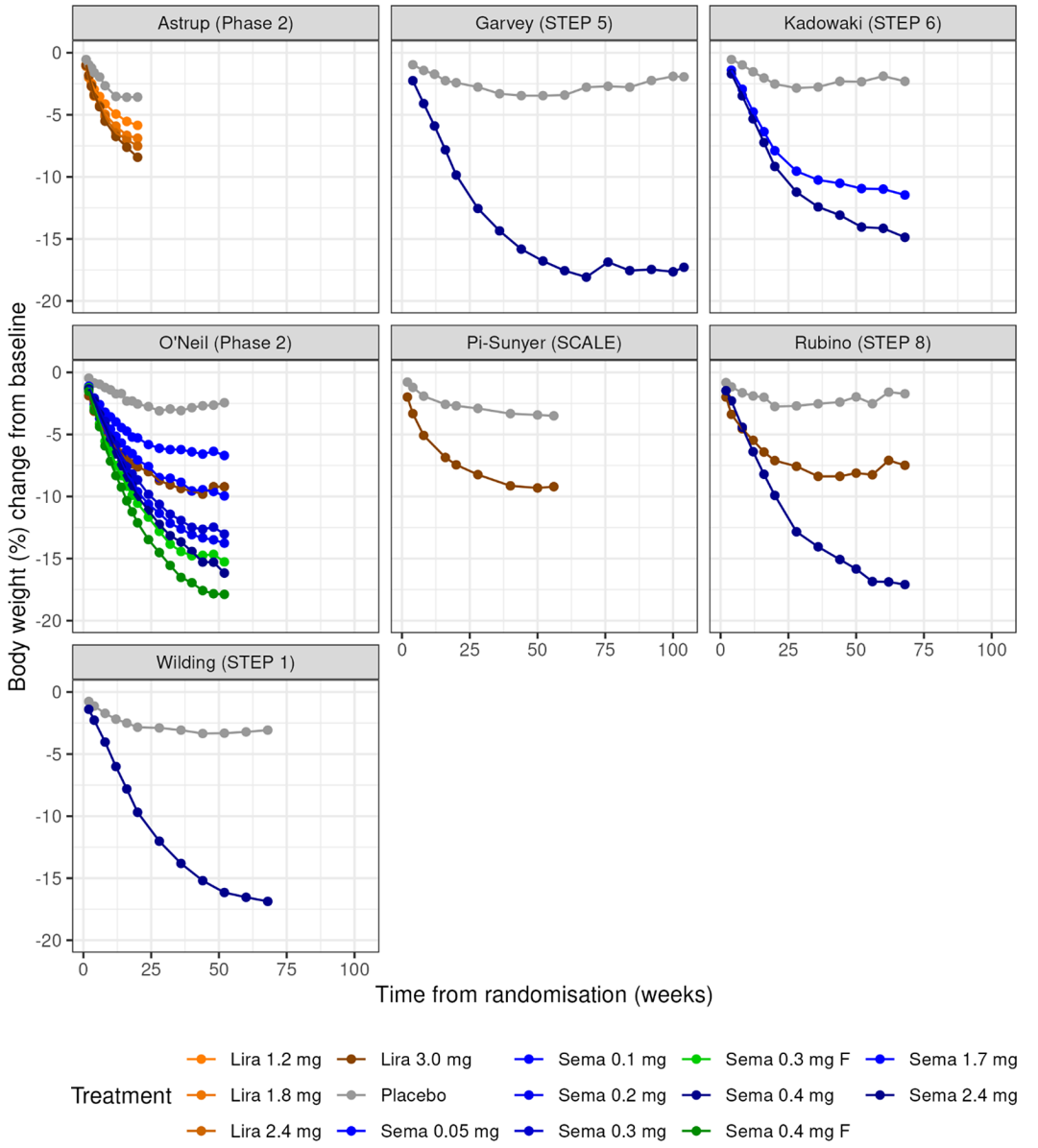


Figure 2. *Plots of body weight change over time for each of the studies in the obesity dataset. Studies are labelled by first author (study name). Panels show the observed mean percentage change from baseline in body weight over time among participants in the full analysis population during the on-treatment observation period. Semaglutide dose levels in O'neil (Phase 2) were standardized to weekly equivalents by multiplying daily doses with 7 to match the once-weekly doses used in phase 3. The treatment labelled 'F' represents doses, where semaglutide was dose-escalated in every second to target doses 0.30 mg and 0.40 mg, respectively.*

heterogeneity between studies increases as consequence of a poor performing dose–response function. Embedding an E_{max} function on both time–course parameters (E_{max}T and kT) to model the dose–response had convergence problems and is therefore not reported. In Figure 3, the VPCs (Visual Predictive Check) are displayed, and in Figure S6 in the Supplementary Material, deviance–deviance plot is shown for relevant dose–response models compared to the UME. Additionally, the individual

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

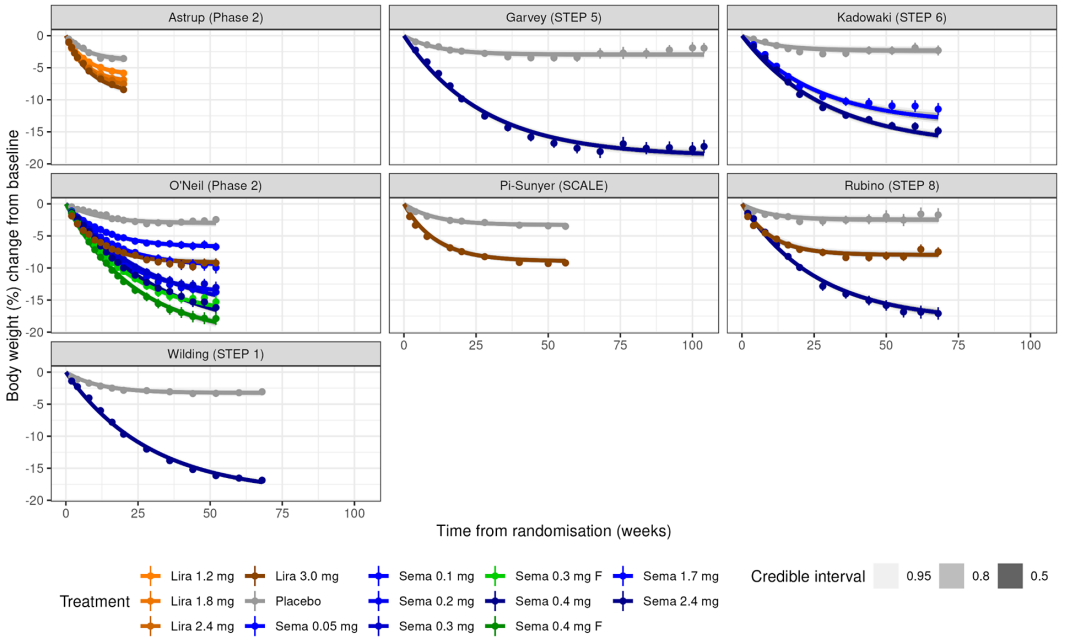


Figure 3. *Plots of body weight change over time for each of the studies in the obesity dataset. Studies are labelled by first author (study name). Panels show the observed mean percentage change from baseline in body weight over time among participants in the full analysis population during the on-treatment observation period. Lines are model posterior means and credible intervals.*

studies are provided in Figures S7–S12 in the Supplementary Material, providing more insight into model performance. This suggests that DT-MBNMA adequately described the summary-level data for all study arms.

ITC between different liraglutide and semaglutide doses that are approved for the treatment of obesity[12, 13] are presented in Table 7. With the MBNMA approach, we can rank agents at any given dose on both structural parameters in the exponential time–course function, but we are interested in treatment nadir (i.e., long-term weight loss), and hence dose-dependent effects on the E_{max}T parameter are used for the ITC. The relative effect of semaglutide 2.4 mg vs. liraglutide 3 mg is in line with study levels results from Rubino (STEP 8), showing that semaglutide 2.4 mg provides a superior weight reduction compared to liraglutide 3 mg.[21] The relative effect of semaglutide 1.7 mg vs. semaglutide 2.4 mg is in line with study levels results from Kadowaki (STEP 6).[20] There is no direct evidence for comparison of semaglutide 1.7 mg vs. liraglutide 3 mg. Both methods (T-MBNMA and NMA) indicate that semaglutide 1.7 mg provides a superior weight reduction compared to liraglutide 3 mg, but the relative treatment effect appears larger with DT-MBNMA method. Here the difference between the two methods is clearer in the sense that the DT-MBNMA uses information from the full treatment network (i.e., phase 2 and phase 3, Figure 1a) to estimate the dose–response relationship for both agents, whereas NMA is informed only by phase 3 data, where the treatment effect of semaglutide 1.7 mg is only observed and reported in the Kadowaki (STEP 6) study. Therefore, the ITC of semaglutide 1.7 mg vs. liraglutide 3.0 mg can be done with greater precision because the treatment comparisons are computed from the agent specific dose–response curves, being summarized in only two drug effect parameters.

Finally, because IPD were available for these illustrative trials, we computed the observed within-arm correlation matrices to assess sensitivity of the results to the correlation specification. This is used in the estimation of drug effect parameters and thus of indirect treatment effects. The use of observed within-arm correlation structure had limited effects on the relative treatment effects under the DT-MBNMA approach. This is in line with results from the simulation study.

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

Table 7. *Indirect treatment comparison (ITC) between different liraglutide and semaglutide doses that are approved for the treatment of obesity.*

Indirect treatment comparison	Estimate	Method	Within-correlation
Semaglutide 1.7 mg vs. liraglutide 3.0 mg	−4.87 [−6.85; −2.75]	DT-MBNMA[a]	Observed
Semaglutide 1.7 mg vs. semaglutide 2.4 mg	2.6 [2.36; 2.99]	DT-MBNMA[a]	Observed
Semaglutide 2.4 mg vs. liraglutide 3.0 mg	−7.56 [−9.77; −5.28]	DT-MBNMA[a]	Observed
Semaglutide 1.7 mg vs. liraglutide 3.0 mg	−6.8 [−8.0; −5.4]	DT-MBNMA[b]	AR–1
Semaglutide 1.7 mg vs. semaglutide 2.4 mg	3.3 [2.5; 4.1]	DT-MBNMA[b]	AR–1
Semaglutide 2.4 mg vs. liraglutide 3.0 mg	−10.1 [−11.3;−8.9]	DT-MBNMA[b]	AR–1
Semaglutide 1.7 mg vs. liraglutide 3.0 mg	−5.1 [−8.0; −2.6]	T-MBNMA[b]	AR–1
Semaglutide 1.7 mg vs. semaglutide 2.4 mg	5.4 [2.7; 7.7]	T-MBNMA[b]	AR–1
Semaglutide 2.4 mg vs. liraglutide 3.0 mg	−10.4 [−12.0; −8.9]	T-MBNMA[b]	AR–1
Semaglutide 1.7 mg vs. liraglutide 3.0 mg	−3.8 [−7.4; −0.72]	NMA (P)[c]	
Semaglutide 1.7 mg vs. semaglutide 2.4 mg	4.0 [1.0; 6.9]	NMA (P)[c]	
Semaglutide 2.4 mg vs. liraglutide 3.0 mg	−7.8 [−10.2; −6.0]	NMA (P)[c]	
Semaglutide 1.7 mg vs. liraglutide 3.0 mg	−4.4 [−7.78; −1.27]	NMA (E)[c]	
Semaglutide 1.7 mg vs. semaglutide 2.4 mg	4.1 [1.3; 7.1]	NMA (E)[c]	
Semaglutide 2.4 mg vs. liraglutide 3.0 mg	−8.5 [−10.8; −6.5]	NMA (E)[c]	
Semaglutide 2.4 mg vs. liraglutide 3.0 mg	−9.4 [12.0; −6.8] (treatment policy) −10.5 [−12.8; −8.1] (hypothetical)	Study results (direct evidence)[d]	
Semaglutide 1.7 mg vs. semaglutide 2.4 mg	3.7 [no test] (treatment policy)	Study results (direct evidence)[d]	

[a]The exponential function $E_{max}T\ (1 - \exp\left(-\exp(kT) \times time\right))$ is specified for the time–course relationship. Treatment contrasts are derived from the semaglutide- and liraglutide-specific log-linear dose–response functions. The within-arm correlation was the observed within treatment arm .

[b]The exponential function $E_{max}T\ (1 - \exp\left(-\exp(kT) \times time\right))$ is specified for the time–course relationship. Treatment contrasts are derived from the semaglutide- and liraglutide-specific $E_{\max}$ dose–response functions or agent-by-dose combinations. The within-arm correlation was the observed approximated with an AR-1 process.

[c]The Standard NMA analysis used two datasets: (P) trial level estimates from the primary estimand specified per protocol and (E) trial level estimates from the end-of-treatment visit as observed, to align with the data used in MBNMA. MBNMA and NMA estimates are presented as posterior mean with 95% credible intervals in brackets.

[d]Trial level results comparing semaglutide 1.7 mg vs. semaglutide 2.4 mg in STEP 6 and semaglutide 1.7 mg vs. semaglutide 2.4 mg in STEP 8.

4. Discussion

4.1. Joint estimation of dose–response and time–course relationship

The purpose of the simulation study was to demonstrate that the drug effect parameters can be estimated jointly with time–course parameters that may be dose-dependent, while respecting randomization. Simulations show that drug effect parameters can be estimated reliably and with low bias. The precision of the drug effect parameters and therefore indirect treatment effects is affected by the quantity and location of observations at a given set of timepoints. These findings are consistent with the results of an extensive simulation study,[23] investigating aspects of the aggregate data for parameter identifiability. We refer to Pedder et al.[23] for a discussion on how T-MBNMA parameters were affected by the quantity and location of observed timepoints across studies. In addition, we have also shown how the observed within-arm correlations can be used in the aggregated data analysis although the information is seldom available in publications.

Convergence failure was not observed in any of the simulated datasets, which can be attributed to the simulation setup. It must be noted that three or more doses are required to be tested in the RCTs to estimate a dose–response relationship (e.g., Figure S2B in the Supplementary Material) and preferably the included RCTs should cover a broad dose-range. For instance, optimal experimental design theory has shown that to estimate parameters ED50 and E_{max} with the highest precision, then a dose should be located at the point of median effective dose and a dose should be located at the point, where the response is plateauing. Therefore, it is expected that in the design of individual dose-finding studies, careful consideration is given to the number and location of doses in the grid, enabling dose–response functions to be fitted for the characterization of the dose–response relationship for a given agent. Proof-of-concept studies typically test a single high dose against placebo and can be included directly in the analysis to provide evidence on maximal efficacy, though they contribute no information for dose–response parameter estimation. Alternatively, such studies may inform hierarchical priors or sensitivity analyses.

Future research should systematically evaluate the framework's performance under dose–response misspecification. While we demonstrate unbiased recovery under correct specification and provide tools to detect misspecification (ELPD, between-study heterogeneity diagnostics), simulations systematically quantifying bias when incorrect functional forms are fitted would complement this work.

The framework accommodates agent-specific parameters with the possibility of combining these into shared parameters or class-effect parameters. Hence, information on dose–response characteristics can be inferred from other agents by assuming varying degrees of similarity and thus improving parameter identifiability. In a reimbursement context, ITCs frequently need to make auxiliary assumptions to have a connected network. For example, windows for analysis timepoints may need to be defined; otherwise, the treatment network will be disconnected. Such windows may differ from the decision problem and/or introduce bias. A parametric time–course component allows the inclusion of all timepoints, and predictions can be made at timepoints relevant for the decision.

We have shown how the NMA framework can be extended using ideas from the MBMA literature to incorporate both dose–response and time–course models in a coherent manner. The framework is used to analyse data from seven RCTs in the treatment of obesity with GLP-1 RAs, covering a wide dose span and end-of-treatment timepoints varying from 20 weeks to 104 weeks. In a standard NMA, phase 2 would have been excluded due to the short treatment duration even though a licensed dose (i.e., liraglutide 3.0 mg) was included in the Astrup 2009 trial.[6] As illustrated here, all available evidence is integrated with the MBNMA method at the expense of additional assumptions on the parametric forms for the time–course relationship. We have used a 'staged' approach for model selection as advocated by Pedder et al.,[3,23] and the selected DT-MBNMA is compared to the T-MBNMAs that do not make any assumptions of the shape of the dose–response. As an alternative to this approach, Bayesian model averaging (BMA) may be an attractive method. BMA is a statistical method used to address model uncertainty in the context of model selection. Rather than selecting a single 'best' model, BMA considers a set of candidate models and accounts for uncertainty by averaging over the predictions of multiple models, weighted by their posterior probabilities. The methodology has recently

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

been employed in dose–response modelling[24] and meta-analysis applications[25] and could be explored for MBNMA.

Joint modelling of dose–response and time–course relationships requires sufficient data across both the dose and time dimensions. Our obesity case study demonstrates a setting where the framework is applicable. However, therapeutic areas with sparse phase 2 evidence, limited dose ranges in published trials, or few intermediate timepoints may lack the data needed for reliable estimation. In such cases, simpler approaches (standard NMA with timepoint windows or separate dose–response and time–course models) may be preferable. Importantly, the framework can also be used to tackle heterogeneous treatment durations and dose selections within phase 3 evidence, which would otherwise require arbitrary timepoint windows or disconnect the network.

4.2. The role of model-based network meta-analysis in drug development

The approach described in this paper can be used in model-informed drug development as a platform for evaluation of indirect comparative efficacy and safety information. For example, to evaluate a proposed phase 3 trial comparing dose X to comparator Y at 52 weeks, we simulate outcomes by drawing from the DT-MBNMA posterior predictive distribution of treatment effects at the planned dose and timepoint. The Bayesian formulation captures both parameter uncertainty (from the dose–response and time–course models) and between-trial heterogeneity, enabling estimation of power or probability of success for the planned design. Such simulations can inform dose selection and sample size requirements by quantifying uncertainty in expected treatment effects for designs not yet tested. While we have focused on estimating indirect treatment effects, the framework also enables prediction of absolute time–course profiles for specific doses and timepoints.

5. Conclusion

We have devised a statistically robust framework that combines dose–response MBNMA and T-MBNMA. The resulting DT-MBNMA framework utilizes parametric models to formalize Bayesian evidence synthesis across trials with between-trial differences in assessment timepoints and dose levels. In addition to being useful for informing decision-making in clinical development, the DT-MBNMA framework is a valuable addition to the HTA toolbox, enabling stronger integration of evidence.

Competing interest statement. The authors declare that they are shareholders in Novo Nordisk A/S.

Author contributions. Conceptualization: A.S., M.B., A.G.-R.; Data curation: A.S., M.B., A.G.-R.; Formal analysis: A.S., M.B., A.G.-R.; Methodology: A.S., M.B., A.G.-R.; Software: A.S.; Writing—original draft: A.S.; Writing—review and editing: A.S., M.B., A.G.-R.

Data availability statement. The program code file and simulated dataset can be found at https://zenodo.org/records/19850316.

Funding statement. The project was internally supported by Novo Nordisk A/S.

References

[1] Pedder H, Dias S, Bennetts M, Boucher M, Welton NJ. Joining the dots: linking disconnected networks of evidence using dose–response model-based network meta-analysis. *Med Decis Mak*. 2021;41(2): 194–208.
[2] Chan P, Peskov K, Song X. Applications of model-based meta-analysis in drug development. *Pharm Res*. 2022;39(8): 1761–1777.
[3] Pedder H, Dias S, Bennetts M, Boucher M, Welton NJ. Modelling time-course relationships with multiple treatments: model-based network meta-analysis for continuous summary outcomes. *Res Synth Methods*. 2019;10(2): 267–286.
[4] Mawdsley D, Bennetts M, Dias S, Boucher M, Welton NJ. Model-based network meta-analysis: a framework for evidence synthesis of clinical trial data. *CPT Pharmacometrics Syst Pharmacol*. 2016;5(8): 393–401.
[5] Abdelaal M, le Roux CW, Docherty NG. Morbidity and mortality associated with obesity. *Ann Transl Med*. 2017;5(7): 161.
[6]]r6 Astrup A, Rossner S, Van Gaal L, et al. Effects of liraglutide in the treatment of obesity: a randomised, double-blind, placebo-controlled study. *Lancet*. 2009;374(9701): 1606–1616.
[7] Higgins JP, Whitehead A. Borrowing strength from external trials in a meta-analysis. *Stat Med*. 1996;15(24): 2733–2749.
[8] Pedder H, Dias S, Boucher M, Bennetts M, Mawdsley D, Welton NJ. Methods to assess evidence consistency in dose–response model based network meta-analysis. *Stat Med*. 2022;41(4): 625–644.

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press

[9] Bergstrand M, Hooker AC, Wallin JE, Karlsson MO. Prediction-corrected visual predictive checks for diagnosing nonlinear mixed-effects models. *AAPS J*. 2011;13(2): 143–151.
[10] Garrett ER. The Bateman function revisited: a critical reevaluation of the quantitative expressions to characterize concentrations in the one compartment body model as a function of time with first-order invasion and first-order elimination. *J Pharmacokinet Biopharm*. 1994;22(2): 103–128.
[11] Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Comput*. 2017;27: 1413–1432.
[12] Novo Nordisk A/S. Saxenda® (liraglutide), US Prescribing Information (PI). 2023.
[13] Novo Nordisk A/S. Wegovy® (semaglutide), US Prescribing Information (PI). 2024.
[14] Pi-Sunyer X, Astrup A, Fujioka K, et al. A randomized, controlled trial of 3.0 mg of liraglutide in weight management. *N Engl J Med*. 2015;373(1): 11–22.
[15] O'Neil PM, Birkenfeld AL, McGowan B, et al. Efficacy and safety of semaglutide compared with liraglutide and placebo for weight loss in patients with obesity: a randomised, double-blind, placebo and active controlled, dose-ranging, phase 2 trial. *Lancet*. 2018;392(10148): 637–649.
[16] Kapitza C, Nosek L, Jensen L, Hartvig H, Jensen CB, Flint A. Semaglutide, a once-weekly human GLP-1 analog, does not reduce the bioavailability of the combined oral contraceptive, ethinylestradiol/levonorgestrel. *J Clin Pharmacol*. 2015;55(5): 497–504.
[17] Strathe A, Horn DB, Larsen MS, et al. A model-based approach to predict individual weight loss with semaglutide in people with overweight or obesity. *Diabetes Obes Metab*. 2023;25(11): 3171–3180.
[18] Wilding JPH, Batterham RL, Calanna S, et al. Once-weekly semaglutide in adults with overweight or obesity. *N Engl J Med*. 2021;384(11): 989–1002.
[19] Garvey WT, Batterham RL, Bhatta M, et al. Two-year effects of semaglutide in adults with overweight or obesity: the STEP 5 trial. *Nat Med*. 2022;28(10): 2083–2091.
[20] Kadowaki T, Isendahl J, Khalid U, et al. Semaglutide once a week in adults with overweight or obesity, with or without type 2 diabetes in an east Asian population (STEP 6): a randomised, double-blind, double-dummy, placebo-controlled, phase 3a trial. *Lancet Diabetes Endocrinol*. 2022;10(3): 193–206.
[21] Rubino DM, Greenway FL, Khalid U, et al. Effect of weekly subcutaneous semaglutide vs daily liraglutide on body weight in adults with overweight or obesity without diabetes: the STEP 8 randomized clinical trial. *JAMA*. 2022;327(2): 138–150.
[22] Brkner P-C. brms: an R package for Bayesian multilevel models using Stan. *J Stat Softw*. 2017;80: 1–28.
[23] Pedder H, Boucher M, Dias S, Bennetts M, Welton NJ. Performance of model-based network meta-analysis (MBNMA) of time-course relationships: a simulation study. *Res Synth Methods*. 2020;11(5): 678–697.
[24] Gould AL. BMA-mod: a Bayesian model averaging strategy for determining dose–response relationships in the presence of model uncertainty. *Biom J*. 2019;61(5): 1141–1159.
[25] Gronau QF, Heck DW, Berkhout SW, Haaf JM, Wagenmakers E-J. A primer on Bayesian model-averaged meta-analysis. *Adv Methods Pract Psychol Sci*. 2021;4(3): 1–11.

Cite this article: Strathe A, Bøg M, Gorst-Rasmussen A. Model-based network meta-analysis: Joint estimation of dose–response and time–course relationships. *Research Synthesis Methods*. 2026;00: 1–20. https://doi.org/10.1017/rsm.2026.10098

https://doi.org/10.1017/rsm.2026.10098 Published online by Cambridge University Press