Incorporating dose effects in network meta-analysis

Jennifer A. Watt, MD, PhD[1,2] jennifer.watt@utoronto.ca
Cinzia Del Giovane, PhD[3] cinzia.delgiovane@biham.unibe.ch
Dan Jackson, PhD[4] daniel.jackson1@astrazeneca.com
Rebecca M. Turner, PhD[5] becky.turner@ucl.ac.uk
Andrea C. Tricco, PhD[6] andrea.tricco@unityhealth.to
Dimitris Mavridis, PhD[7] dmavridi@uoi.gr
Sharon E. Straus, MD, MSc[1,2,5] sharon.straus@utoronto.ca
Areti Angeliki Veroniki, PhD[1,8*] areti-angeliki.veroniki@unityhealth.to

* Corresponding Author
Areti Angeliki Veroniki, MSc, PhD
209 Victoria Street, East Building, Toronto, Ontario
M5B 1T8, Canada
Phone: 416-564-5015; Fax: 416-564-5735; Email: areti-angeliki.veroniki@unityhealth.to

[1]Knowledge Translation Program, Li Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria Street, Toronto, Ontario, Canada, M5B 1W8
[2]Division of Geriatric Medicine, Department of Medicine, University of Toronto, 190 Elizabeth Street, R. Fraser Elliott Building, 3-805, Toronto, Ontario, Canada, M5G 2C4
[3]Institute of Primary Health Care (BIHAM), University of Bern, Bern, Switzerland
[4]Statistical Innovation Group, AstraZeneca, 1 Francis Crick Avenue, Cambridge, United Kingdom, CB2 0AA
[5]MRC Clinical Trials Unit, University College London, 90 High Holborn, London, United Kingdom, WC1V 6LJ
[6]Institute of Health Policy, Management and Evaluation; University of Toronto, Toronto, Ontario, Canada
[7]Department of Primary Education, School of Education, University of Ioannina, Ioannina, Greece
[8]Institute of Reproductive and Developmental Biology, Department of Surgery & Cancer, Faculty of Medicine, Imperial College, London, United Kingdom

Keywords: network meta-analysis, systematic review, dose, antipsychotics, cholinesterase inhibitors, cerebrovascular event, nausea, headache

Word count: 4418 (manuscript); 1 table, 7 figures, 4 boxes, 1 supplement file

The final published version of this article may be found at:
https://www.bmj.com/content/376/bmj-2021-067003

KEY MESSAGES BOX

1. Systematic reviews with network meta-analysis (NMA) that ignore potential dose effects may limit the applicability and validity of review findings.
2. Hierarchical random effects NMA models incorporating dose effects assume dose level consistency and that dose effects are equal (model 1), separate (model 2), or exchangeable (model 3). These NMA models do not make assumptions about the shape of dose response relationships.
3. While researchers should first consider clinical and pharmacological factors when selecting the most appropriate NMA model for their clinical question, statistical and methodological considerations such as between study and between dose heterogeneity, consistency across treatment and dose effects, and model fit are also important.
4. Clinicians and other knowledge users should appraise the applicability and validity of NMA modelling assumptions, including explanations of the model selection process and biological plausibility for incorporating (or not incorporating) dose effects.

CONTRIBUTORS AND SOURCES

Dr. Jennifer Watt is a geriatrician with experience in applying network meta-analysis models. She contributed to the study design, prepared datasets, completed analyses, drafted the initial manuscript, and integrated co-author feedback. Dr. Cinzia Del Giovane, Dr. Rebecca Turner, Dr. Dan Jackson, and Dr. Dimitris Mavridis are statisticians with expertise in developing and applying network meta-analysis models. Dr. Andrea Tricco is a methodologist and Dr. Sharon Straus is a geriatrician – both have expertise and experience in conducting systematic reviews with network meta-analysis to support clinical and policy decision-making. They

contributed to the study design and provided manuscript feedback. Dr. Areti-Angeliki Veroniki is a statistician with network meta-analysis expertise. She is co-Convenor of the Cochrane Statistical Methods Group. She conceived the study idea, developed model codes, completed analyses, drafted the initial manuscript, and provided manuscript feedback. We used data from two published systematic reviews and network meta-analyses.[1 2]

STANDFIRST

Systematic reviews with network meta-analysis (NMA) that ignore potential dose effects may limit the applicability and validity of review findings; here, we help content experts (e.g., clinicians), methodologists, and statisticians better understand how to incorporate dose effects in network meta-analysis by (1) describing three network meta-analysis models that make different clinical and statistical assumptions about how to model dose effects, (2) illustrating the importance of dose effects in understanding the potential risk of harm in people with dementia from cerebrovascular events associated with atypical antipsychotic use (i.e., quetiapine, olanzapine, and risperidone) and nausea and headache associated with cholinesterase inhibitor use (i.e., donepezil, galantamine, and rivastigmine), and (3) discussing important considerations when choosing between different network meta-analysis models incorporating dose effects.

INTRODUCTION

In contrast to pairwise meta-analysis, which directly compares one treatment's efficacy or safety to another based on head-to-head data, network meta-analysis (NMA) simultaneously compares and ranks multiple treatments that are either directly compared through head-to-head data, indirectly compared through a common treatment comparator, or both (i.e., a mixed treatment comparison composed of direct and indirect evidence).[3] If a researcher wants to compare the efficacy or safety of multiple treatments, NMA can better answer this question than pairwise meta-analysis. The ability of NMA to simultaneously compare the efficacy and safety of multiple treatments has led to a sharp rise in the number of published NMAs and research to improve their methodological rigor.[4]

NMAs improve decision making by filling knowledge gaps where no head-to-head comparative treatment data exist, but an absence of NMA results concerning treatment dose effects could limit their applicability and validity. For example, although it is helpful to know that donepezil, galantamine, and rivastigmine (medications used to improve symptoms of Alzheimer disease) are associated with an increased risk of nausea, clinicians could better support tailored decision making if they know which medication doses are associated with this risk.[2] A lack of methodological guidance for researchers on how to incorporate treatment dose effects into systematic reviews with NMAs is contributing to this critical omission. Our objective is to present three hierarchical NMA models that researchers can implement to incorporate dose effects into systematic reviews with NMA, even in the absence of prior knowledge of how to model a dose-response relationship; give practical guidance on how to conduct these analyses; provide empirical examples so readers can appreciate the importance of modelling dose effects; describe considerations for evaluating the appropriateness of NMA models incorporating dose effects; discuss considerations in appraising the applicability and validity of systematic reviews with NMA incorporating dose effects; and highlight

challenges, limitations, and future research directions related to selection of NMA models incorporating dose effects. Our empirical examples describe the dose effect association between (1) atypical antipsychotic use and risk of cerebrovascular events in people with dementia and (2) cholinesterase inhibitor use and risk of nausea or headache in people with Alzheimer disease, but the NMA models incorporating dose effects that we describe could be applied to examples in any medical discipline.[1 2]

MODIFYING HIERARCHICAL NMA MODELS TO INCORPORATE DOSE EFFECTS

In the standard NMA model, consistency is assumed, random treatment effects are modelled, and effect estimates (e.g., odds ratios, mean differences) are derived on the treatment level; dose effects are not explicitly modelled.[3] In this current paper, we show how this hierarchical NMA model can be modified to incorporate dose effects.

Let us consider a hypothetical network of five treatments T=(a, b, c, d, e) and 11 different doses indexed with t_i^T, i=1,…,11. In Figure 1*a*, treatment 'a' is the network reference node, which is a treatment with a single or no dose (e.g., placebo), and other nodes represent treatments 'b', 'c', 'd', and 'e'; in Figure 1*b*, we see that treatments are composed of doses. Here, we present three hierarchical random effects NMA models incorporating dose effects (see also Supplement File 1), which differ based on: (1) if they assume consistency on the treatment level (i.e., between direct and indirect comparisons); (2) the number of variance components; (3) if they account for the relationship between dose and parent treatment; and (4) whether effect estimates are derived on the treatment level, dose level, or both (Table 1).[5] None of these NMA models assume a parametric dose response relationship.

Table 1. Properties of three hierarchical network meta-analysis dose effects models

Characteristic	Equal dose effects (model 1)	Separate dose effects (model 2)	Exchangeable dose effects (model 3)

Accounts for within study variation	Yes	Yes	Yes
Accounts for between study variation at the dose level by incorporating random dose effects	Yes	Yes	Yes
Accounts for between dose within treatments variation using a variance component	No	No	Yes
Assumes consistency on treatment level	Yes	No	No^{+}
Assumes consistency on dose level	Yes*	Yes	Yes
Exchangeability of dose-effects within treatments/includes between dose variance component	No‡	No^{++}	Yes**
Accounts for the treatment dose relationship	Yes††	No	Yes
Produces effect estimates on the treatment level	Yes	No	Yes
Produces effect estimates on the dose level	No	Yes	Yes

+ Consistency is assumed on the dose level, and treatment effects are assumed to be exchangeable within doses. This does not imply treatment-effect consistency in the conventional sense.

* Consistency is assumed on the dose-level, and since all doses within the same treatment are assumed to be equally effective, consistency is also assumed at the treatment level.

‡ Average dose effects are identical within treatments, a stronger assumption than exchangeable dose effects within treatments.

++ Doses are considered unrelated with respect to their parent treatment. Model 2 is equivalent to the conventional consistency model for network meta-analysis, where each treatment-dose combination is treated as a different group.

** Doses are related and exchangeable within their parent treatment.

†† Model 1 accounts for the treatment dose relationship in a simple way; whereby all average dose effects are the same in the same parent treatment.

There are three main sources of variation in hierarchical random effects NMA dose effects models (Table 1 and Figure 2): within study, between study, and between dose within treatments. The first level of variation is within studies (i.e., the variability across study participants), which is modelled in a conventional way whereby each study has its own study specific baseline.[3] The second level of variation is between studies: the variability in true effects across studies within each treatment dose comparison.[6] In contrast to the standard NMA model, where between study variation is modelled at the treatment level, hierarchical NMA dose effects models incorporate between study variation at the dose level.[3] In a random effects model, each study specific true effect size is part of a distribution of all true effect sizes and the variance of this distribution represents the between study variance. There

is also a third level of variation: the between dose variation within treatments. This refers to the variability of dose effects within each treatment category, assuming that each dose corresponds to a specific treatment category. All three hierarchical NMA dose effects models incorporate within and between study variation in the same way; however, only the exchangeable NMA dose effects model (model 3) incorporates variance components for all three potential sources of variation.

Equal dose effects (model 1)

The simplest NMA model incorporates equal average dose effects (Table 1 and Figure 3a). This approach can only be considered for research questions targeted at assessing treatment effects, as it assumes that different doses of the same treatment are fixed and equally effective or safe within the same treatment group. This NMA model may include studies with data on multiple doses for the same treatment, but the dose effects are fixed and equal to the broader treatment effect. Data from study arms where the relative effects are assumed equal to zero contribute to the between study variance estimation. An equal dose effects NMA model accounts for within study and between study variation, assumes consistency on the treatment and dose levels, and produces effect estimates (e.g., log-odds ratio) on the treatment level.

Separate dose effects (model 2)

This NMA model incorporates separate average dose effects (Table 1 and Figure 3b). It is appropriate for research questions assessing the effects of specific treatment doses, as it accounts for different dose effects. This NMA model assumes that average dose effects are unrelated with respect to their parent treatment and each other, and each node in the network is a separate treatment dose; therefore, the treatment dose relationship is not considered. The separate dose effects NMA model accounts for within study and between study variation, assumes consistency on the dose level only, and produces effect estimates on the dose level.

Exchangeable dose effects (model 3)

This NMA model assumes that the average dose effects are related and exchangeable within their parent treatment (also known as 'exchangeable sub-nodes'; Table 1 and Figure 3c).[7] This NMA model accounts for the treatment dose relationship, distinguishes between different treatment doses, and assumes that average dose effects within the same treatment come from a common distribution. It accounts for within study, between study, and between dose variation within treatments using variance components; assumes consistency on the dose level only; and produces effect estimates on both the treatment and dose levels. Because this model does not require additional assumptions about how to model the shape of the dose response relationship (like models 1 and 2); accounts for the treatment dose relationship (like model 1); distinguishes between different treatment doses (like model 2); explicitly models within study, between study, and between dose variation within treatments using variance components; and produces effect estimates on both the treatment and dose levels, the exchangeable dose effects NMA model is a preferred NMA model for understanding different treatment doses if statistical and methodological considerations are valid (e.g., dose level consistency and transitivity) (Boxes 1 and 2). When the between dose variance is estimated as zero, this model simplifies to the equal dose effects NMA model (i.e., model 1).

Box 1. Considerations in choosing network meta-analysis models incorporating treatment and dose effects

1. Anticipated clinical significance of treatment and dose effects (i.e., network meta-analysis results should incorporate clinically relevant dose effects)
2. Between study and between dose heterogeneity
3. Appropriateness of assuming transitivity and consistency on the treatment level, dose level, or both
4. Model fit and parsimony
5. Network geometry, connectedness (i.e., avoidance of disconnected network components), and sparsity

Box 2. Advantages to implementing hierarchical network meta-analysis models incorporating random dose effects within treatments (model 3)

1. Considers the treatment dose relationship

2. Does not make any parametric assumptions about potential dose response relationships
3. Facilitates borrowing of strength within treatment classes when different doses are available
4. Allows for the inclusion of studies comparing only multiple doses of the same treatment
5. Facilitates the simultaneous identification of the best treatment and dose
6. Can increase power compared to carrying out several independent subgroup analyses or extreme splitting approaches (i.e., model 2)

ILLUSTRATIVE EXAMPLES

We illustrate the aforementioned NMA models with three empirical examples, which are presented below.[1 2] For each example, we present: (1) network plots; (2) transitivity tables; (3) model fit statistics (i.e., deviance information criterion [DIC]); (4) between study and between dose heterogeneity estimates; (5) global (i.e., design-by-treatment interaction model) and local (i.e., loop-specific approach) inconsistency estimates at the treatment and dose levels; (6) outcomes as medians with 95% credible intervals (CrIs) and 95% prediction intervals (PrIs); and (7) rankings according to surface under the cumulative ranking curve (SUCRA) values (i.e., 100% indicates the best performing treatment and 0% indicates the worst).[5 8-10] We summarized SUCRA values for each outcome across models in a rank heat plot.[11] We performed analyses in OpenBUGS (model fit and estimation methods are described in Supplement File 2; OpenBUGS model code is available in Supplement File 3; and all study data, transitivity tables, model fit statistics, heterogeneity estimates, inconsistency plots, treatment and dose level outcomes, and treatment and dose rankings are found in Supplement Tables 1 to 12; and Figures 1, 2, and 3).[12]

Atypical antipsychotics

Dataset

Antipsychotics are prescribed to people with dementia for treating neuropsychiatric symptoms (e.g., aggression), but they are associated with potential harms in this patient

population, including an increased risk of cerebrovascular events.[1 13] Our example dataset is a subset of data describing the risk of cerebrovascular events associated with atypical antipsychotic use (i.e., quetiapine, olanzapine, or risperidone) in people with dementia, which was published in a systematic review and NMA describing the comparative safety of pharmacologic interventions for treating neuropsychiatric symptoms in people with dementia (Supplement Table 1).[1] Here, we include only those randomized trials that reported a target or average total daily treatment dose. We categorized treatment doses based on average total daily dose, where reported; otherwise, we categorized doses using target total daily dose. We categorized atypical antipsychotic doses as per ranges proposed by Maust et al: low dose quetiapine (<125mg/day), medium dose quetiapine (125mg/day to 200mg/day), high dose quetiapine (>200mg/day), low dose olanzapine (<5mg/day), medium dose olanzapine (5mg/day to <7.5mg/day), high dose olanzapine (≥7.5mg/day), low dose risperidone (≤1mg/day), medium dose risperidone (>1mg/day to 2mg/day), and high dose risperidone (>2mg/day).[14]

Results: Cerebrovascular Events

We included 10 studies (3,079 patients), four treatments, and seven treatment doses in our hierarchical NMA models incorporating treatment and dose effects (Figure 4a). There were differences in dementia types and study duration across treatment and dose comparisons (Supplemental Tables 2 and 3). Small between study heterogeneity was evident in the network, which did not importantly change across models (Supplement Tables 4a-d). Model fit was similar across models. We did not identify any global or local inconsistency at the treatment or dose levels (Supplement Figures 1a and 1b). These results suggest that researchers could implement model 1, 2, or 3, depending on their clinical or policy question.

In model 1, olanzapine (OR 3.18, 95% CrI 1.12 to 9.52, 95% PrI 0.97 to 10.75) and risperidone (OR 3.59, 95% CrI 1.71 to 8.03, 95% PrI 1.42 to 9.43) were associated with greater odds of cerebrovascular events compared to placebo. In models 2 and 3, medium dose olanzapine, low dose risperidone, and medium dose risperidone were associated with greater odds of cerebrovascular events compared to placebo (Figure 5 and Supplement Tables 4a-c). With respect to treatment rankings (i.e., SUCRA values), model 1 suggested that quetiapine was the safest and risperidone was the most harmful. With respect to treatment dose rankings, model 2 suggested that low dose olanzapine was the safest; low and medium dose risperidone were the most harmful. Model 3 suggested that low and medium dose quetiapine were the safest; whereas, low and medium dose risperidone were the most harmful (Figure 6a and Supplement Table 4d). Our results suggest that both low dose olanzapine and low and medium dose quetiapine are the safest treatment options for people with dementia because they are not associated with increased odds of cerebrovascular events.

Cholinesterase inhibitors

Datasets

Cholinesterase inhibitors (i.e., donepezil, galantamine, and rivastigmine) are prescribed to people with dementia to slow cognitive decline. However, they are associated with potential harms, including nausea and headache.[2] Our example datasets are subsets of data describing the risk of nausea and headache associated with cholinesterase inhibitor use in people with Alzheimer disease, which were published in a systematic review and NMA describing the comparative effectiveness and safety of cognitive enhancers in people with Alzheimer disease (Supplement Tables 5 and 6).[2] Here, we include only those randomized trials that reported a target or average total daily treatment dose. We categorized treatment doses based on average total daily dose, where reported; otherwise, we categorized treatment doses based

upon target total daily dose. We categorized cholinesterase inhibitor doses as per ranges proposed by Lee et al: low dose donepezil (≤5mg/day), high dose donepezil (>5mg/day), low dose galantamine (<16mg/day), high dose galantamine (≥16mg/day), low dose rivastigmine (<6mg/day), and high dose rivastigmine (≥6mg/day).[15]

Results: Nausea

We included 41 studies (10,604 patients), four treatments, and seven treatment doses in our hierarchical NMA models describing the association between cholinesterase inhibitor use and nausea (Figure 4b). Study and participant characteristics were similar across treatment and dose comparisons (Supplement Tables 7 and 8). Moderate between study heterogeneity was evident in model 1 (0.20, 95% CrI 0.06 to 0.49), which decreased substantially in models 2 and 3 (Supplement Tables 9a-c). Model 1 (DIC=157) fit the data better than models 2 (DIC=167) and 3 (DIC=165). Although no inconsistent network loops were evident at the treatment level, inconsistency was identified at the dose level for the loop involving placebo, low dose donepezil, and high dose galantamine (Supplement Figures 2a and 2b). Given the presence of one inconsistent network loop at the dose level, researchers could cautiously proceed with implementing models 1, 2, or 3; however, they could consider an alternative approach (Box 3).[3] Lower between study heterogeneity in models 2 and 3 than model 1 suggests that treatment dose explains part of the between study heterogeneity. If researchers proceed with implementing NMA models that assume consistency on the dose level because of important clinical considerations, they should implement model 2 or 3, depending on whether they are interested in dose effects only (i.e., model 2) or treatment and dose effects (i.e., model 3).

In model 1, donepezil (OR 1.72, 95% CrI 1.24 to 2.45, 95% PrI 0.65 to 4.70), galantamine (OR 2.98, 95% CrI 2.05 to 4.31, 95% PrI 1.09 to 8.12), and rivastigmine (OR 3.78, 95% CrI

2.61 to 5.59, 95% PrI 1.4 to 10.44) were associated with greater odds of nausea compared to placebo. In model 2, high dose rivastigmine was associated with greater odds of nausea than all other treatments; and high dose galantamine was associated with greater odds of nausea than high dose donepezil, low dose donepezil, and placebo. In model 3, high dose rivastigmine was associated with greater odds of nausea compared to all treatments except high dose galantamine; and high dose galantamine was associated with greater odds of nausea than high dose donepezil, low dose donepezil, and placebo (Figure 7a and Supplement Tables 9a-c). With respect to treatment rankings, model 1 suggested that placebo was the safest and rivastigmine was the most harmful treatment. Models 2 and 3 suggested there was a dose response across treatment doses (i.e., high treatment doses had the least favorable treatment dose profiles; Figure 6b and Supplement Table 9d). Our results suggest that high dose galantamine and high dose rivastigmine are associated with increased odds of nausea in people with Alzheimer disease and that low rather than high cholinesterase inhibitor doses are associated with more favorable nausea risk profiles.

Results: Headache

We included 31 studies (8,589 patients), four treatments, and seven treatment doses in our hierarchical NMA models describing the association between cholinesterase inhibitor use and headache (Figure 4c). Study and participant characteristics were similar across treatment comparisons, but there were differences across dose comparisons with regards to study duration (Supplement Tables 10 and 11). Between study heterogeneity was greatest in model 1 (0.28, 95% CrI 0.07 to 0.76). DICs across models were similar. There was no evidence of inconsistency at the treatment or dose levels (Supplement Figures 3a and 3b). These findings suggest that researchers should implement model 2 or 3 because of the lower estimated between study heterogeneity in these models compared to model 1, depending on whether

interest lies in deriving dose effects only (i.e., model 2) or both treatment and dose effects (i.e., model 3).

In model 1, only rivastigmine was associated with increased odds of headache compared to placebo (OR 2.19, 95% CrI 1.35 to 3.62, 95% PrI 0.65 to 7.57). In model 2, high dose rivastigmine was associated with increased odds of headache compared to placebo, high dose donepezil, and low dose rivastigmine. In model 3, only high dose rivastigmine was associated with increased odds of headache compared to placebo (Figure 7b and Supplement Table 12a-c). With respect to treatment ranking, model 1 suggested that placebo was the safest and rivastigmine was the most harmful treatment. Models 2 and 3 suggested there was a dose response across treatment doses (i.e., high treatment doses had the least favorable treatment dose profiles; Figure 6c and Supplement Table 12d). Our results suggest that high dose rivastigmine is associated with increased odds of nausea in people with Alzheimer disease and that low rather than high cholinesterase inhibitor doses are associated with more favorable headache risk profiles.

DISCUSSION

Clinical importance of modelling both treatment and dose effects

It is important to use NMA models that reflect real life clinical experiences; if studies incorporate clinically relevant treatment doses, then researchers should use NMA models incorporating dose effects so that results are responsive to the needs of decision makers unless there are methodological or statistical considerations that will jeopardize NMA conclusions (Box 1). For this reason, the equal dose effects model (model 1) is only recommended when it is plausible to assume that any dose effects are very small or absent because model 1 ignores possible differences in dose effects within treatments (Box 2 and Figure 3a). Like model 3, model 2 incorporates both treatment and dose effects, but model 2

ignores potential treatment dose relationships; derives only dose effects; and does not explicitly model between dose variation within treatments using variance components. Model 3 is a highly appropriate model in the presence of different dose effects for helping decision makers to understand the comparative efficacy or safety of multiple treatments and doses simultaneously. Hierarchical NMA models can also be extended to cases where describing the effects of treatment formulations (e.g., oral, intravenous) and potential effect modifiers (i.e., meta-regression) is important. Further, these NMA models could be modified to incorporate a parametric dose response. .Our examples demonstrate both treatment and dose effects, which provide decision makers with important information beyond what was previously available in published medical literature.[1 13 16] First, our results showed that risperidone and medium dose olanzapine were associated with increased odds of cerebrovascular events, which may prompt clinicians to prescribe quetiapine or low dose olanzapine to avoid this feared adverse event. Second, we demonstrated a potential treatment and dose response relationship for the outcome of nausea across cholinesterase inhibitors – low dose donepezil was the best tolerated and high dose rivastigmine was the worst tolerated. However, decision makers need to cautiously interpret these findings since we detected local inconsistency in this NMA model. Lastly, if we had modelled only treatment effects, we would have assumed all doses of rivastigmine were associated with increased risk of headache; by incorporating dose effects, we found that this increased risk was associated with high dose rivastigmine only.

Dose effects as a source of heterogeneity

NMA models should reflect our real-life clinical understanding of treatment doses: we assume that there is a treatment dose relationship (i.e., doses of one treatment are more similar than are doses of another treatment) and how we model heterogeneity should reflect this understanding (Box 1). Further, if the estimated between study variation is sensitive to

model choice, then reviewers can investigate with subgroup, sensitivity, or meta-regression analyses to understand if dose variability is an effect modifier or if participant characteristics vary by treatment dose (Table 1). For example, in our empirical examples involving cholinesterase inhibitors, the equal dose effects model (model 1) increased estimated between study heterogeneity compared to the separate (model 2) and exchangeable (model 3) dose effects models.

Appropriateness of assuming transitivity and consistency on the treatment level, dose level, or both

Transitivity implies that effect modifiers are balanced across NMA treatment and dose comparisons; consistency is the statistical quantification of transitivity. Researchers should evaluate these assumptions on each level that they are assumed (i.e., transitivity and consistency assumptions must be assessed on both the treatment and dose levels if researchers apply model 1). In addition to intransitivity or inconsistency related to dose effects, inconsistency may also be due to an imbalance in the distribution of other effect modifiers (e.g., participant age, sex, dementia severity). We did not identify any global or local inconsistency on the treatment level in our examples. On the dose level, we identified one inconsistent network loop in our example where we described the association between cholinesterase inhibitor use and risk of nausea; dose effects estimated from direct evidence were significantly different from dose effects estimated from indirect evidence in the closed network loop incorporating placebo, low dose donepezil, and high dose galantamine.[5] Where inconsistency or intransitivity is identified on the dose level, it may not be appropriate to apply a model assuming consistency on the dose level and researchers should consider alternative approaches (Box 3).[3 17] Researchers need to explore a number of factors (e.g., between-study variance, transitivity, consistency, model fit statistics) before choosing between models (Box 1). Fitting multiple models could improve understanding of the data set

and interpretation of results. Readers and peer reviewers of manuscripts reporting NMAs incorporating treatment and dose effects should also consider these factors when appraising the applicability and validity of systematic reviews with NMA (Box 4).

Box 3. Alternative knowledge synthesis approaches when it is potentially inappropriate to assume consistency on the dose level in NMA models

1. Apply a model that assumes consistency on the treatment level only (i.e., model proposed by Dias et al.)[3]
2. Incorporate random inconsistency effects in the dose effects model
3. Explore inconsistency and intransitivity through meta-regression or subgroup analyses
4. Apply pairwise meta-analysis models only
5. Narratively synthesize systematic review findings without performing meta-analysis

Box 4. Considerations in appraising the applicability and validity of systematic reviews with NMA incorporating treatment and/or dose effects

1. Is the biological plausibility of incorporating dose effects explained?
2. If authors decide to incorporate dose effects, have they included all clinically relevant treatment doses?
3. If authors chose a NMA model incorporating dose effects, do they provide a valid rationale for their model selection process?
4. If NMA models incorporating dose effects assume a dose response relationship, have authors justified how they chose to model this dose response relationship?
5. If dose effects are not incorporated in NMA models, have authors explained why they made this decision (e.g., network sparsity, dose level inconsistency, poor model fit, no biological plausibility, not relevant to the research question)?

Alternative approaches for incorporating dose effects

Alternative approaches to modelling dose effects in NMAs have been suggested.[7 18-20] Del Giovane et al. proposed a number of other hierarchical NMA models incorporating dose effects.[7] Similar to model 3, reviewers could apply a random dose effects NMA model without assuming consistency on the dose level; however, this model can only be implemented in the case where there are no multi-arm studies.[7] Del Giovane et al. also proposed that adjacent treatment doses could be modelled as more similar than non-adjacent doses with a random walk process or it could be assumed that there is a monotonic dose response relationship (e.g., higher doses are likely to be more beneficial for clinical outcomes).[7] These alternative hierarchical NMA models incorporating dose effects require

that researchers make additional modelling assumptions, which should be carefully considered *a priori* by a multidisciplinary team (e.g., content experts, methodologists, and statisticians). Owen et al. proposed a hierarchical NMA model that assumed a monotonic but nonparametric dose response between nodes representing different doses of the same treatment.[20] Owen et al. implemented ordering constraints (i.e., assumed that higher doses would be associated with the same or greater clinical benefit).[20] Thorlund et al. implemented a network meta-regression model that assumed a linear dose response on the log-odds scale and incorporated a three-level categorical covariate for doses at (1) half each drug's "common" dose, (2) each drug's "common dose", or (3) double each drug's "common" dose.[19] In this model, assumptions must be made about what each drug's "common dose" is, which can vary by study population. Mawdsley et al. proposed a model-based NMA framework that facilitates estimation and prediction of dose effects for multiple treatments within a drug class across a range of doses (including those for which study data are not available), using plausible physiological dose response models.[18]

Challenges and limitations of applying NMA models incorporating treatment and dose effects

There are challenges and potential limitations to applying NMA models incorporating dose effects. First, studies that do not report treatment dosing information cannot be included in NMA models incorporating treatment and dose effects. Second, performing NMAs that assume equal average dose effects (model 1) may increase precision of treatment effects, but there are potential trade-offs: (1) greater heterogeneity and inconsistency if there are clinically meaningful dose effects that are not included in the NMA model; and (2) NMA outputs that are potentially less meaningful for decision makers, especially if it is believed that dose effects are clinically important. In contrast, "splitting" of treatment nodes into smaller dose-based sub-nodes may decrease precision in effect estimates because there are

fewer studies informing each NMA dose comparison (model 2), but (1) heterogeneity and inconsistency may decrease because the effects of dose on heterogeneity are explicitly modelled and (2) NMA outputs will potentially be more meaningful for decision makers. Third, "splitting" of nodes to incorporate dose effects may create treatment doses with zero events or disconnected networks. Fourth, decisions about how to model dose response relationships in NMA models can be complicated, which is why we present three NMA models incorporating dose effects that do not require prior knowledge of this dose response relationship; however, if researchers have confidence in how to model the dose response relationship for treatments under study then alternative models can be considered, as proposed by Del Giovane and others.[7 18 20] Fifth, given that studies reporting dose effects may have more than two arms and comparison-adjusted funnel plots assume independence between effect estimates in multi-arm studies, researchers assessing for publication bias can instead implement a selection model (e.g., Copas model) and present funnel plots for each direct treatment comparison.[21 22] Most direct treatment comparisons in our NMA models were informed by fewer than 10 studies so we could not evaluate for publication bias. Lastly, we implemented NMA models in a Bayesian framework, which may be less familiar to some researchers, but NMA models incorporating dose effects could alternatively be implemented in a frequentist framework. A Bayesian framework offers several advantages compared to a frequentist framework, including modelling flexibility, a simpler way to derive ranking statistics associated with treatment and dose effects, the ability to implement informative priors to estimate between-study variance, and a more intuitive interpretation of results for decision makers.

CONCLUSION

Having the ability to incorporate both treatment and dose effects is important for researchers whose goal is to produce relevant and clinically meaningful NMA results for decision

makers. However, implementing NMA models incorporating treatment and dose effects is complex and requires the skills of a multidisciplinary team (e.g., clinicians, methodologists, and statisticians). As we have highlighted, clinical and pharmacological considerations should be considered first, but statistical and methodological considerations are also important. Further, different approaches and decisions about network structure may generate important variations in results so, when possible, decisions concerning NMA model assumptions should be made *a priori*. Future research to guide selection of NMA models incorporating dose effects will be critical to developing a consensus-based approach and advancing knowledge synthesis methods incorporating NMA.

FIGURES

Figure 1. Fictional example with network nodes representing treatments (a) and doses (b)

Figure 2. Graphical representation of sources of variance in dose effects models

Figure 3. Graphical representation of networks according to how dose effects are incorporated into network meta-analysis models (equal [a], separate [b], and exchangeable [c] dose effects).

Figure 4. Network diagrams depicting network connectedness of treatments and treatment-doses for three illustrative examples: (a) cerebrovascular events, (b) nausea, and (c) headache. Thickness of solid lines is proportional to the number of studies included in the group comparison, and node size is proportional to the number of patients included in the underlying group. Dashed oval lines group doses of the same treatment.

Figure 5. Forest plot of odds ratios (OR; 95% credible intervals [CrI]) describing the association between atypical antipsychotic (i.e., olanzapine, quetiapine, and risperidone) treatment doses and odds of cerebrovascular event compared to placebo. Blue triangles represent the summary dose effects derived from model 2 and red circles represent the summary dose effects derived from model 3. There are four treatments and seven treatment doses.
Abbreviations: medium dose olanzapine (OLA-M), medium dose quetiapine (QUET-M), medium dose risperidone (RIS-M), low dose olanzapine (OLA-L), low dose quetiapine (QUET-L), low dose risperidone (RIS-L).

Figure 6. Rank-heat plots for the outcomes of (a) cerebrovascular events, (b) nausea, and (c) headache across treatment and treatment doses. Each model corresponds to a separate ring. Sectors are coloured according to surface under the cumulative ranking curve (SUCRA) values as per the transformation of three colours red (0%), yellow (50%), and green (100%). Circles from outside in refer to: 1st, equal dose effects (model 1); 2nd, separate dose effects (model 2); 3rd, exchangeable dose effects (model 3).
Abbreviations: high dose donepezil (DON-H), high dose galantamine (GAL-H), medium dose olanzapine (OLA-M), medium dose quetiapine (QUET-M), medium dose risperidone (RIS-M), high dose rivastigmine (RIV-H), low dose donepezil (DON-L), low dose galantamine (GAL-L), low dose olanzapine (OLA-L), low dose quetiapine (QUET-L), low dose risperidone (RIS-L), low dose rivastigmine (RIV-L), and placebo (PLA).

Figure 7. Forest plot of odds ratios (OR; 95% credible intervals [CrI]) describing the association between cholinesterase inhibitor (i.e., donepezil, galantamine, and rivastigmine) treatment-doses and

odds of (a) nausea and (b) headache compared with placebo. Blue triangles represent the summary dose effects derived from model 2 and red circles represent the summary dose effects derived from model 3. There are four treatments and seven treatment doses.
Abbreviations: high dose donepezil (DON-H), high dose galantamine (GAL-H), high dose rivastigmine (RIV-H), low dose donepezil (DON-L), low dose galantamine (GAL-L), low dose rivastigmine (RIV-L).

DATA SHARING

The full dataset and statistical code are available in the supplement file.

DISSEMINATION

We will disseminate our results to relevant knowledge user groups (e.g., patients, caregivers, healthcare managers, and clinicians).

ETHICS APPROVAL

Not applicable.

TRANSPARENCY STATEMENT

JAW affirms that this manuscript is an honest, accurate, and transparent account of the study being reported and that no important aspects of the study have been omitted.

ROLE OF THE FUNDING SOURCE

There was no funding for this study. ACT is supported by a Tier 2 Canada Research Chair in Knowledge Synthesis. SES is supported by a Tier 1 Canada Research Chair in Knowledge Translation. RT is supported by the UK Medical Research Council (grant number MC_UU_12023/24).

CONTRIBUTORS

All study authors contributed to the conception and design of this study. JAW and AAV conducted data analyses. JAW and AAV drafted the first version of the manuscript. All authors contributed to the manuscript's revision and interpretation of findings. JAW is the guarantor of this article.

DECLARATION OF INTERESTS

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf. JAW, CDG, RMT, ACT, DM, SES, and AAV declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work. DJ declares that he is employed by AstraZeneca.

AUTHOR LICENSE

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees.

REFERENCES

1. Watt JA, Goodarzi Z, Veroniki AA, et al. Safety of pharmacologic interventions for neuropsychiatric symptoms in dementia: a systematic review and network meta-analysis. *BMC Geriatr* 2020;20(1):212. doi: 10.1186/s12877-020-01607-7 [published Online First: 2020/06/18]
2. Tricco AC, Ashoor HM, Soobiah C, et al. Comparative Effectiveness and Safety of Cognitive Enhancers for Treating Alzheimer's Disease: Systematic Review and Network Metaanalysis. *J Am Geriatr Soc* 2018;66(1):170-78. doi: 10.1111/jgs.15069
3. Dias S, Sutton AJ, Ades AE, et al. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making* 2013;33(5):607-17. doi: 10.1177/0272989X12458724 [published Online First: 2012/10/30]
4. Zarin W, Veroniki AA, Nincic V, et al. Characteristics and knowledge synthesis approach for 456 network meta-analyses: a scoping review. *BMC Medicine* 2017;15(3)
5. Veroniki AA, Vasiliadis HS, Higgins JP, et al. Evaluation of inconsistency in networks of interventions. *International journal of epidemiology* 2013;42(1):332-45.
6. Lu G, Ades A. Modeling between-trial variance structure in mixed treatment comparisons. *Biostatistics* 2009;10(4):792-805. doi: 10.1093/biostatistics/kxp032 [published Online First: 2009/08/19]
7. Del Giovane C, Vacchi L, Mavridis D, et al. Network meta-analysis models to account for variability in treatment definitions: application to dose effects. *Stat Med* 2013;32(1):25-39. doi: 10.1002/sim.5512
8. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *Journal of clinical epidemiology* 2011;64(2):163-71. doi: 10.1016/j.jclinepi.2010.03.016
9. Higgins JP, Jackson D, Barrett JK, et al. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Res Synth Methods* 2012;3:98-110. doi: 10.1002/jrsm.1044
10. van Aert RCM, Schmid CH, Svensson D, et al. Study specific prediction intervals for random-effects meta-analysis: A tutorial: Prediction intervals in meta-analysis. *Res Synth Methods* 2021 doi: 10.1002/jrsm.1490 [published Online First: 2021/05/04]
11. Veroniki AA, Straus SE, Rucker G, et al. Is providing uncertainty intervals in treatment ranking helpful in a network meta-analysis? *Journal of clinical epidemiology* 2018;100:122-29. doi: 10.1016/j.jclinepi.2018.02.009 [published Online First: 2018/02/13]
12. Lunn D, Spiegelhalter D, Thomas A, et al. The BUGS project: Evolution, critique and future directions. *Statistics in medicine* 2009;28(25):3049-67. doi: 10.1002/sim.3680 [published Online First: 2009/07/25]
13. Schneider LS, Dagerman K, Insel PS. Efficacy and adverse effects of atypical antipsychotics for dementia: meta-analysis of randomized, placebo-controlled trials. *Am J Geriatr Psychiatry* 2006;14(3):191-210. doi: 10.1097/01.JGP.0000200589.01396.6d [published Online First: 2006/03/01]
14. Maust DT, Kim HM, Seyfried LS, et al. Antipsychotics, other psychotropics, and the risk of death in patients with dementia: number needed to harm. *JAMA Psychiatry* 2015;72(5):438-45. doi: 10.1001/jamapsychiatry.2014.3018 [published Online First: 2015/03/19]
15. Lee PE, Hsiung G-YR, Seitz D, et al. Cholinesterase Inhibitors. *BC Medical Journal* 2011;53(8):404-08.

16. Yunusa I, Alsumali A, Garba AE, et al. Assessment of Reported Comparative Effectiveness and Safety of Atypical Antipsychotics in the Treatment of Behavioral and Psychological Symptoms of Dementia: A Network Meta-analysis. *JAMA Netw Open* 2019;2(3):e190828. doi: 10.1001/jamanetworkopen.2019.0828 [published Online First: 2019/03/23]
17. Jackson D, Barrett JK, Rice S, et al. A design-by-treatment interaction model for network meta-analysis with random inconsistency effects. *Statistics in medicine* 2014;33(21):3639-54. doi: 10.1002/sim.6188 [published Online First: 2014/04/30]
18. Mawdsley D, Bennetts M, Dias S, et al. Model-Based Network Meta-Analysis: A Framework for Evidence Synthesis of Clinical Trial Data. *CPT: pharmacometrics & systems pharmacology* 2016;5(8):393-401. doi: 10.1002/psp4.12091
19. Thorlund K, Mills EJ, Wu P, et al. Comparative efficacy of triptans for the abortive treatment of migraine: a multiple treatment comparison meta-analysis. *Cephalalgia* 2014;34(4):258-67. doi: 10.1177/0333102413508661 [published Online First: 2013/10/11]
20. Owen RK, Tincello DG, Keith RA. Network meta-analysis: development of a three-level hierarchical modeling approach incorporating dose-related constraints. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2015;18(1):116-26. doi: 10.1016/j.jval.2014.10.006
21. Mavridis D, Welton NJ, Sutton A, et al. A selection model for accounting for publication bias in a full network meta-analysis. *Statistics in medicine* 2014;33(30):5399-412. doi: 10.1002/sim.6321 [published Online First: 2014/10/16]
22. Chaimani A, Higgins JP, Mavridis D, et al. Graphical tools for network meta-analysis in STATA. *PLoS One* 2013;8(10):e76654. doi: 10.1371/journal.pone.0076654 [published Online First: 2013/10/08]