

Dose–effect meta-analysis for psychopharmacological interventions using randomised data

Tasnim Hamza ^{1,2}, Toshi A Furukawa ³, Nicola Orsini,⁴ Andrea Cipriani ⁵,
Georgia Salanti ¹

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/ebmental-2021-300278>).

¹Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland

²Graduate School for Health Sciences, University of Bern, Bern, Switzerland

³Department of Health Promotion and Human Behavior, and Department of Clinical Epidemiology, Graduate School of Medicine/School of Public Health, Kyoto University, Kyoto, Japan

⁴Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

⁵Department of Psychiatry, University of Oxford, Oxford, UK

Correspondence to

Tasnim Hamza, University of Bern Institute of Social and Preventive Medicine, Bern, Switzerland; tasnim.hamza@ispm.unibe.ch

Received 18 May 2021

Accepted 22 June 2021

Published Online First

18 January 2022

ABSTRACT

Objective The current practice in meta-analysis of the effects of psychopharmacological interventions ignores the administered dose or restricts the analysis in a dose range. This may introduce unnecessary uncertainty and heterogeneity. Methods have been developed to integrate the dose–effect models in meta-analysis.

Methods We describe the two-stage and the one-stage models to conduct a dose–effect meta-analysis using common or random effects methods. We illustrate the methods on a dataset of selective serotonin reuptake inhibitor antidepressants. The dataset comprises 60 randomised controlled trials. The dose–effect is measured on an odds ratio scale and is modelled using restricted cubic splines to detect departure from linearity.

Results The estimated summary curve indicates that the probability of response increases up to 30 mg/day of fluoxetine-equivalent which results in reaching 50% probability to respond. Beyond 40 mg/day, no further increase in the response is observed. The one-stage model includes all studies, resulting in slightly less uncertainty than the two-stage model where only part of the data is analysed.

Conclusions The dose–effect meta-analysis enables clinicians to understand how the effect of a drug changes as a function of its dose. Such analysis should be conducted in practice using the one-stage model that incorporates evidence from all available studies.

INTRODUCTION

Synthesis of evidence provided by randomised controlled trials (RCTs) is commonly used to develop clinical guidelines and make reimbursement decision for pharmacological interventions. While the dose of a drug is of central importance, meta-analyses that examine their efficacy and safety often focus on comparing only agents or classes of drugs, ignoring potential variability due to different doses. As different dose schedules may result in considerable heterogeneity in efficacy and safety, one common approach is to restrict the database at certain dose range (e.g., the therapeutic dose), discard all studies outside that range and then examine the role of dose in a subgroup analysis for the lowest and the highest dose categories.¹ This approach fails, however, to synthesise the whole relevant evidence. Alternatively, researchers might opt to perform many meta-analyses, each restricted to studies that examine a particular drug-dose combination. This will inevitably result in many underpowered meta-analyses.

In this paper, we present a recently developed evidence synthesis method of a dose–effect meta-analysis (DE-MA) approach that offers a middle ground between ‘lumping’ all doses together into a single meta-analysis and ‘splitting’ them to many dose-specific meta-analyses. In DE-MA, we model the changes in the drug effect along the range of all studied dosages. There are two common approaches to conduct DE-MA: two-stage and one-stage models. In the two-stage model, the dose–effect curve is estimated within each study and then synthesised across studies.^{2,3} These two steps are performed simultaneously in the one-stage model.⁴

We first provide the statistical explanations of the two models, and then illustrate the models by using a collection of RCTs examining the efficacy of selective serotonin reuptake inhibitors (SSRI) antidepressants.⁵

The analysis is implemented in R⁶ and is made available along with dataset and the results on GitHub (<https://github.com/htx-r/Dose-effect-MA-EBMH-article->).

METHODS

In this section, we describe the two-stage DE-MA model with summarised data. Then we present briefly the one-stage model. Finally, we discuss other issues related to this topic, namely; statistical testing of dose–effect coefficients and how to assess heterogeneity and make predictions. The models which are illustrated here to conduct DE-MA have been implemented in various software packages, for example, the *drmeta* command (in Stata⁷) and the *dosresmeta* package⁸ (in R).⁶

Dose–effect shape within a study

Let us consider the case of an RCT where several doses are examined (one dose per arm) denoted by $dose_j$ where the index j enumerates the dose levels starting with zero. The outcome is measured in each arm on an additive scale (e.g., a mean, a log-odds). The dose–effect model within a study associates the change in the outcome (ie, the treatment effect) to the change in the dose. Let us assume a trial like the one presented in [table 1](#) that has a placebo arm, a dichotomous outcome and the changes in the outcome are measured using the odds ratio (logOR) of each dose level j relevant to a reference dose $dose_0$. Using the placebo arm as a reference (at dose ($dose_0 = 0$)), and assuming a linear association between logOR and dose, the dose–effect model is

$$\log OR_j = \beta(dose_j - dose_0)$$



© Author(s) (or their employer(s)) 2022. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Hamza T, Furukawa TA, Orsini N, et al. *Evid Based Ment Health* 2022;**25**:1–6.

Table 1 We present the data of Feighner *et al* study on the observed five dose levels. The data consist of the observed dose, the number of responses, the total number of participants, the odds ratio (OR) and its 95% CI, log transformations of OR and its standard error (SE)

Level (j)	fluoxetine-equivalent dose (mg/day)	Response	Total	OR (95% CI)	logOR	SE(logOR)
0	0	42	129	1.00 (ref)	0 (ref)	
1	10	61	131	1.81 (1.09 to 2.99)	0.59	0.26
2	20	61	130	1.83 (1.11 to 3.03)	0.61	0.26
3	40	80	131	3.25 (1.95 to 5.41)	1.18	0.26
4	60	73	129	2.70 (1.63 to 4.48)	0.99	0.26

The estimated coefficient β shows how much an increase in the dose will impact on the change in logOR.

Typically, the referent dose $dose_0$ is assigned to the zero or the minimal dose to make interpretation easier. The doses are centred around the referent dose so the relationship quantifies the change in relative effects. However, this centralisation induces correlation between the logORs in each study (as they are all estimated relative to the outcome of the ($dose_0$)). Such correlations should be estimated and accounted for using the Longnecker and Greenland method.^{2,9}

In practice, multiple changes in the dose–effect shape are expected so that the linear model is not often a realistic assumption. More flexible models are needed to account for those changes¹⁰ such as restricted cubic spline (RCS). RCS is a piecewise function; the dose spectrum is split into intervals (using some changepoints, called knots) and in each interval a cubic polynomial is fitted.¹¹ Restrictions in the estimation of the polynomial coefficients are then imposed to ensure that they are connected and forming a smooth function which is linear in the two tails. The location and the number of those knots determine the shape of the RCS: the locations indicate intervals where changes in the shape might occur, and the number reflects how many such changes are anticipated. In general, setting k knots creates a RCS model with $k - 1$ regression coefficients. For identifiability, the minimum number of knots is three and the dose–effect shape is:

$$\log OR_j = \beta_1 \{dose_j - dose_0\} + \beta_2 \{f_1(dose_j) - f_1(dose_0)\}$$

This function is a combination of linear and non linear (f_1) transformations.¹¹

Of note, a two-stage approach requires that the study examines at least three dose-level data including the referent level and that enables estimating the two regression coefficients in the linear (β_1) and spline (nonlinear, β_2) parts of the equation.

Any type of function could be used in the dose–effect association. For study indicator i , the general form of the dose–effect model can be written:

$$\log OR_{ij} = f(dose_{ij}; dose_{i0}; \beta_{ip})$$

The term β_{ip} refers to the p dose–effect parameter and f denotes the dose–effect shape.

Synthesis of dose–effect shapes across studies

Consider that we have fit the RCS model in k studies and we have obtained k sets of estimates (β_{i1}, β_{i2}). Each pair of coefficients represents the shape of the dose–effect within each study. Now, we synthesise the shapes across studies by combining their coefficients. We may set a common underlying coefficient for all studies, for example, $\beta_{i1} = B_1$ and $\beta_{i2} = B_2$ (common-effect model). Alternatively, the underlying study-specific coefficients can be assigned a two-dimensional normal distribution with mean (B_1, B_2) and a variance–covariance matrix to reflect the

heterogeneity across the studies (random-effects model). In the general case, the dose–effect shape f involving p coefficients β_{ip} which are similarly synthesised using a multivariate normal distribution.

What we describe above is the two-stage approach: the dose–effect curves are estimated within each study and then synthesised across studies in two separate steps. This requires each study to report non-referent doses at least as many as the number of the dose–effect coefficients. Otherwise, the coefficients will be non-identifiable and the study should be excluded from the analysis. For example, to estimate a dose–effect quadratic shape or a RCS with three knots, two coefficients need to be estimated and hence each study needs to report at least two logORs (which means at least three dose levels). Studies that report less dose levels, shall be excluded from the synthesis.

In the one-stage approach, within and across study estimation of the shape are performed simultaneously.⁴ This allows for borrowing information across studies and the study-specific coefficients can be estimated even if the study itself does not report the required number of doses. This means that, with the one-stage approach, we can include in the synthesis studies that report only one logOR (two dose levels) even if we want to estimate RCS.

There are different ways to present the results from the DE-MAs. The dose–effect shape as a function of any dose can be presented in graphical or tabular form by plugging-in the dose values and the estimated coefficients in the assumed function (see figures 1 and 2). Another useful presentation of the results could be to show absolute estimates of the outcome, such as estimates of probability for efficacy at any given dose, see figure 3. This can be done in two simple steps. First, we estimate the absolute probability of the response at the reference dose (e.g., zero) and then we combine this with the estimated relative treatment effect at each dose (e.g., with the estimated logOR) to obtain the absolute outcome (e.g., the probability to respond at an active dose level).

Statistical testing of the dose–effect shape

The hypothesis of no dose–effect association, that is $H_0 : B = 0$ where B is a vector of all regression coefficients, can be tested by computing a Wald statistic based on estimated regression coefficients (and their estimated variances/covariances). A p-value is then derived with reference to a chi-squared distribution (χ^2) with degrees of freedom equal to the number of regression coefficients involved in the null hypothesis.

Alternatively, we may compute the z statistic to test each dose–effect coefficient, B_p under the hypothesis $H_0 : B_p = 0$. Testing the coefficient of the spline term will indicate whether a linear function is sufficient to describe the data.

As with every statistical test, test results should be interpreted with cautious and considering common fallacies and

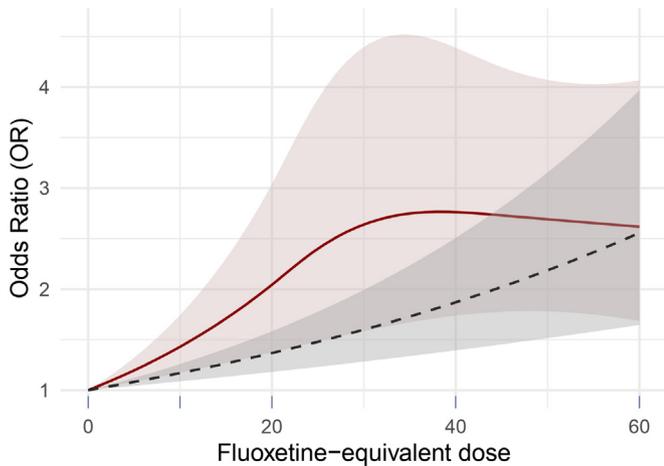


Figure 1 The estimated dose–effect curves of citalopram in Feighner *et al* study. The fluoxetine-equivalent doses are presented versus the odds ratio with two different dose–effect shapes; the linear model in grey (dashed) and the restricted cubic spline (with knots at 20.0, 23.6 and 44.4) in red (solid). The 95% confidence bands are shaded around each curve.

misinterpretations of the p-value.¹² Furthermore, confidence limits, typically 95%, for the unknown summary dose–effect shape can be estimated from the model for any sensible value of the dose.

Heterogeneity

Heterogeneity in the study-specific coefficients β_{ip} introduces heterogeneity in the relative treatment effects, which is what we will call heterogeneity from now on. It is a function of the dose and can be measured by the variance partition coefficient (VPC).⁴ The VPC is a study-specific and dose-specific which shows the percentage of heterogeneity out of the total variability specific to the study. VPC can be computed for each non-referent dose in each study. An average of the study-specific VPCs by dose level could be seen as a dose-specific I^2 . It is useful to plot the

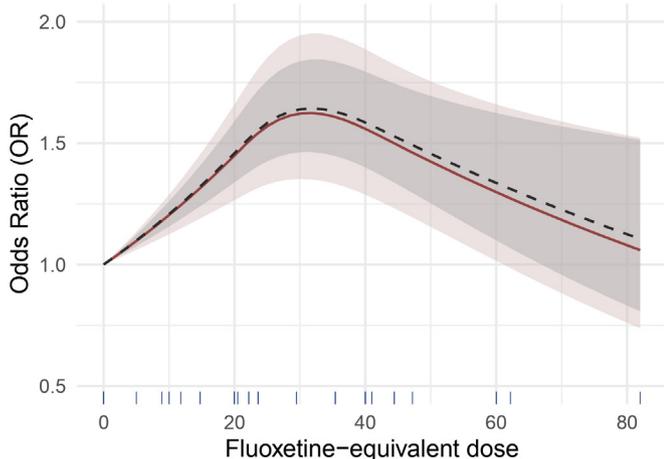


Figure 2 Dose-effect curves for selective serotonin reuptake inhibitors. These curves are estimated using the restricted cubic spline function where knots are set at doses 20.0, 23.6 and 44.4 mg/day. For data synthesis, we apply a one-stage (grey, solid) and two-stage (red, dashed) approaches. The 95% confidence bands are shaded around each curve. SSRI, selective serotonin reuptake inhibitor.

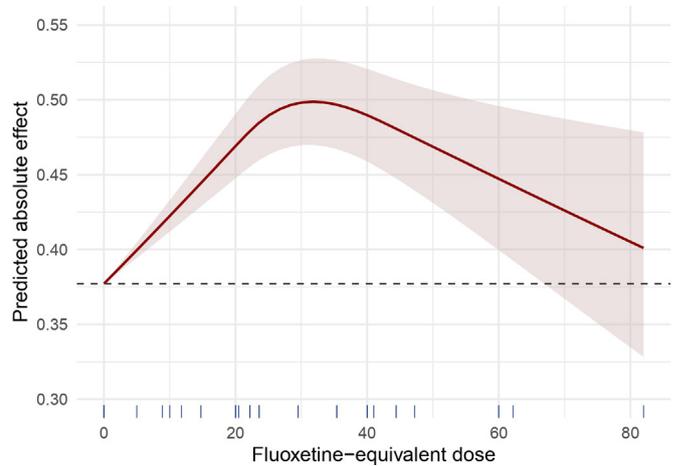


Figure 3 The synthesised dose–effect curves across studies of SSRI. The fluoxetine-equivalent doses are presented versus the predicted absolute effect. The dose–effect function is the restricted cubic spline (with knots at 20.0, 23.6 and 44.4). The solid line represents the mean absolute effect and the shaded area is its 95% confidence bands. The dashed (horizontal) line represents the placebo absolute effect at 37.7%. SSRI, selective serotonin reuptake inhibitor.

study-specific VPCs (as %) against the dose levels to gauge the level of heterogeneity.

RESULTS

We illustrate the models by re-analysing a dataset about the role of dose in the efficacy of SSRIs. Drug-specific doses are converted into fluoxetine-equivalents (mg/day) using a validated formula.⁵ The outcome is response to treatment defined as 50% reduction in symptoms. The data include 60 RCTs, which recruited 15 174 participants in 145 different dose arms (see online supplemental appendix figure 1, 2 and table 1).

Dose–effect model within a study

To exemplify the process, we consider the study by Feighner *et al*.¹³ Table 1 presents the data at the five examined dose arms. The four logORs are estimated as the odds of each non-referent category (10, 20, 40, 60 mg/day) relative to the odds in the referent dose (Placebo, 0 mg/day). The study-specific estimated logORs and their SEs can be used to fit a linear dose–effect model.

A log linear trend is then estimated based on the aggregate data presented by Feighner *et al* (figure 1).¹³ The Greenland and Longnecker method is used to back estimate the covariance of these four empirical logORs used as dependent variable of the linear dose–effect model.

The linear dose–effect coefficient is estimated at 0.0156 (95% CI 0.0083 to 0.0230) on the log scale. The OR at dose 10 to be $exp(0.0156 \times 10) = 1.169$ which means OR increases by 17% for a 10-unit increase in dose.

Biologically, it is quite unrealistic to assume a constant effect of fluoxetine-equivalents on the relative odds of the outcome. We expect the shape to increase up to a dose level and then flatten out. The exact value of the dose, at which the dose–effect model is levelling out, is unknown. And it would be good to specify a dose–effect model that is able to capture this plausible mechanism.

For this reason, we use a RCS function, rather than a linear function, for fluoxetine-equivalents. RCSs are generated using three knots at 20, 23.6 and 44.4 dose levels which represent the

10%, 50% and 90% percentiles, of the observed non-zero dose distribution. A Wald-test indicates large incompatibility between this study and the hypothesis of a linear function ($z = -2.13$, $p = 0.033$). Figure 1 indicates a large positive dose–effect up to 30 mg/day of fluoxetine-equivalents and no increase in the effect beyond that value.

The fact that the shape is estimated from just a single study results in a large uncertainty around the RCS curve.

Synthesis of dose–effect shapes across studies

We first synthesise the dose–effect coefficients from all studies assuming a random-effects two-stage model. For RCS in the two-stage model, only 17 studies can be synthesised (those with at least three dose levels). The results are depicted in figure 2. The estimated linear coefficient at $B_1 = 0.0186$ (95% CI 0.0118 to 0.0253) and the spline coefficient is $B_2 = -0.0628$ (95% CI -0.0876 to -0.0379).

The random-effects one-stage model can include all 60 studies. The estimated linear and spline coefficients are very close to those from the two-stage model ($B_1 = 0.0189$ (95% CI 0.0146 to 0.0232) and $B_2 = -0.0621$ (95% CI -0.0814 to -0.0428)) which is also shown in the agreement of the two shapes in figure 2. The important difference between the results from the two approaches is that the confidence bands are tighter from the one-stage due to including double as many studies as the two-stage approach does.

In figure 3, we show the probability of response as a function of the dose as estimated from the meta-analysis. After meta-analysing all placebo arms, the probability of response to placebo is estimated at 37.7% (dashed line in figure 3). Then, increase of the dose up to 30 mg/day of fluoxetine-equivalent results in 50% probability to respond. Beyond 40 mg/day, the probability of response flattens out.

For the two-stage and the one-stage models, the statistical hypothesis can be rejected with estimated p-values less than 0.001 for both the linear and spline coefficients. This can be seen as a statistical evidence that the linear model hypothesis is rejected, and the RCS is preferable with both the linear and the spline part. The hypothesis of no dose–effect association is not also accepted ($p\text{-value} < 0.001$).

Figure 4 shows the variance partition component along with the observed doses. At dose 20 mg/day, the total variability that is attributed solely to heterogeneity ranges between 4% and 40%, which is considered to be moderate. Overall, the majority of VPC values does not exceed 60%.

DISCUSSION

Researchers can conduct a DE-MA by following two steps. The first step is to estimate a dose–effect curve within each study. The second step is to synthesise those curves across studies. These two steps can be performed either separately (two-stage model)^{2,3} or simultaneously (one-stage model).⁴ In this article, we detail these two models, alongside considerations for statistical testing of the dose–effect parameters, estimation of heterogeneity and presentation of the results. We use the presented models to re-analyse RCT data comparing various SSRIs in terms of response .

We describe the models for a dichotomous outcome and the effect size we used as odds ratio. However, the model can be adapted easily to other measures like risk ratio and hazard ratio. Likewise, the model can be employed with other data types such as continuous outcome with (standardised) mean differences.¹⁴

Recently, two extensions of the presented models have been introduced in the literature. The one-stage and two-stage models have been extended to a Bayesian setting¹⁵ to take advantage of its great flexibility. One of these advantages is to implement the exact binomial distribution for binary data, instead of the

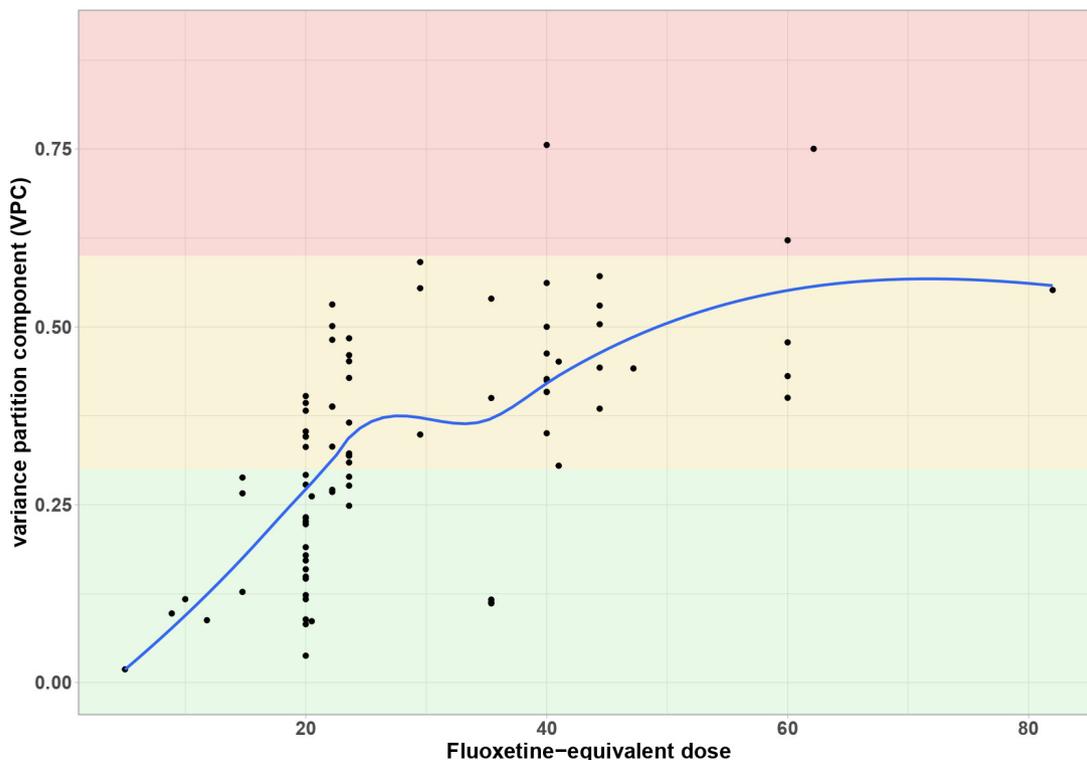


Figure 4 The variance partition component of each observed dose (non-referent doses in each study) presented in circles; each circle represents a study. The fitted line is LOWESS curve.

approximate normal distribution for the relative treatment effect in the frequentist settings. The assumption of a normal distribution can be hard to meet when the sample size is small as shown in recent simulations.¹⁵ The dose–effect model has been also extended to network meta-analysis which allows for modelling the dose–effect relationship simultaneously to more than two agents.^{16 17}

Researchers should be careful when they report the findings of DE-MA and follow the existing reporting guidelines. Xu *et al* proposed a checklist with 33 reporting items for such analysis.¹⁸ The majority of these items (27) come from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement after some modifications.¹⁹ The other six items are added from Meta-analyses Of Observational Studies in Epidemiology checklist to cover key considerations of observational studies.²⁰ They used the proposed checklist to assess quality of reporting in the published DE-MAs. They found that while reporting in the introduction and results was on average good, further improvements are required in reporting methods. Xu and colleagues also studied the association between reporting quality and study characteristics. They observed that studies including more authors or methodologist have a better reporting quality. They conclude that while the quality of reporting has improved over the years, further refinement in the reporting checklists is required.

The main challenge in DE-MA is how to define the dose–effect shape. The shape selection can be guided by previous studies (such as dose-finding studies), clinical experience and biological plausibility informed by pharmacodynamic and pharmacokinetic studies. Additional evidence could be provided by considering the goodness of fitness measures of various shapes²¹ or via graphical inspection of the data. Yet, the RCS model has sufficient flexibility to capture different shapes. In our case study, using only three knots was sufficient to capture the expected drug behaviour SSRIs while requires only three dose levels to be reported in at least one study. This makes RCS an attractive choice for the majority of analyses.¹⁸ However, the number and location of knots should be chosen carefully based on the anticipated drug behaviour and the clinical knowledge.

Researchers may encounter additional challenges if observational studies are synthesised instead of RCTs as it was the case in this paper. First, defining the dependent and independent variables in observational studies could be difficult. For example, if we want to evaluate the association between the alcohol consumption and the use of tobacco, the shape will depend on whether alcohol is set as a dependent or independent variable. Second, categorisation of non-pharmacological exposures (such as environmental exposure, diet and so on), which are often the focus of observational studies, is often difficult. There might be open-ended categories to which assignment of a specific dose is not obvious (e.g., smoking two packages per day and above) and exposure categories might be differently defined across studies.^{22 23} These challenges could induce additional uncertainty in the analysis. In such cases, sensitivity analysis is recommended to investigate the robustness of the DE-MA results.

In conclusion, the DE-MA enables clinicians to understand how the effect of a drug changes as a function of its dose. Such analysis should be conducted in practice using the one-stage model that incorporates evidence from all available studies.

Twitter Toshi A Furukawa @Toshi_FRKW and Andrea Cipriani @And_Cipriani

Contributors TAF and AC designed the study and collected the data. TH performed the statistical analysis and NO helped in a partial revision of the analysis. TH and GS

drafted the manuscript. All authors performed a critical revision of the manuscript and approved the submitted version.

Funding This work was supported by HTx project for TH and GS. The HTx project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825 162. AC is supported by the National Institute for Health Research (NIHR) Oxford Cognitive Health Clinical Research Facility, by an NIHR Research Professorship (grant RP-2017–08–ST2-006), by the NIHR Oxford and Thames Valley Applied Research Collaboration and by the NIHR Oxford Health Biomedical Research Centre (grant BRC-1215–20005). The views expressed are those of the authors and not necessarily those of the UK National Health Service, the NIHR or the UK Department of Health.

Competing interests TAF reports grants and personal fees from Mitsubishi-Tanabe, personal fees from MSD, grants and personal fees from Shionogi, outside the submitted work. In addition, TAF has a patent 2020-548587 concerning smartphone CBT apps pending, and intellectual properties for Kokoro-app licensed to Mitsubishi-Tanabe. AC has received research and consultancy fees from INCIPIT (Italian Network for Paediatric Trials), CARIPO Foundation and Angelini Pharma.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs

Tasnim Hamza <http://orcid.org/0000-0002-4700-6990>

Toshi A Furukawa <http://orcid.org/0000-0003-2159-3776>

Andrea Cipriani <http://orcid.org/0000-0001-5179-8321>

Georgia Salanti <http://orcid.org/0000-0002-3830-8508>

REFERENCES

- Hansen RA, Moore CG, Dusetzina SB, *et al*. Controlling for drug dose in systematic review and meta-analysis: a case study of the effect of antidepressant dose. *Med Decis Making* 2009;29:91–103.
- Berlin J, Longnecker MP, Greenland S. *Meta-analysis of epidemiologic dose-response data*. Epidemiology Resources Inc, 1992.
- Orsini N, Li R, Wolk A, *et al*. Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. *Am J Epidemiol* 2012;175:66–73.
- Crippa A, Discacciati A, Bottai M, *et al*. One-stage dose-response meta-analysis for aggregated data. *Stat Methods Med Res* 2019;28:1579–96.
- Furukawa TA, Cipriani A, Cowen PJ, *et al*. Optimal dose of selective serotonin reuptake inhibitors, venlafaxine, and mirtazapine in major depression: a systematic review and dose-response meta-analysis. *Lancet Psychiatry* 2019;6:601–9.
- Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019;22:153–60.
- Orsini N. *DRMETA: Stata module for dose-response meta-analysis S458546. statistical software components, Boston College department of economics*, 2018.
- Crippa A, Orsini N. Multivariate dose-response meta-analysis: The dosresmeta R Package. *J Stat Softw* 2016;72:1–15.
- Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol* 1992;135:1301–9.
- Bagnardi V, Zambon A, Quatto P, *et al*. Flexible meta-regression functions for modeling aggregate dose-response data, with an application to alcohol and mortality. *Am J Epidemiol* 2004;159:1077–86.
- Harrell FE. *Regression modelling strategies: with applications to linear models, logistic regression, and survival analysis*. New York: Springer, 2015.
- Greenland S, Senn SJ, Rothman KJ, *et al*. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol* 2016;31:337–50.
- Feighner JP, Overø K. Multicenter, placebo-controlled, fixed-dose study of citalopram in moderate-to-severe depression. *J Clin Psychiatry* 1999;60:824–30.
- Crippa A, Orsini N. Dose-response meta-analysis of differences in means. *BMC Med Res Methodol* 2016;16:91.
- Hamza T, Cipriani A, Furukawa TA. *A Bayesian dose–response meta-analysis model: a simulations study and application. statistical methods in medical research*, 2021.
- Mawdsley D, Bennetts M, Dias S, *et al*. Model-Based network meta-analysis: a framework for evidence synthesis of clinical trial data. *CPT Pharmacometrics Syst Pharmacol* 2016;5:393–401.
- Hamza T, Furukawa TA, Orsini N, *et al*. A dose-effect network meta-analysis model: an application in antidepressants. *arXiv* 2104.05414.

- 18 Xu C, Liu T-Z, Jia P-L, *et al.* Improving the quality of reporting of systematic reviews of dose-response meta-analyses: a cross-sectional survey. *BMC Med Res Methodol* 2018;18:157.
- 19 Moher D, Liberati A, Tetzlaff J, *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
- 20 Stroup DF, Berlin JA, Morton SC, *et al.* Meta-Analysis of observational studies in epidemiology: a proposal for reporting. meta-analysis of observational studies in epidemiology (moose) group. *JAMA* 2000;283:2008–12.
- 21 Discacciati A, Crippa A, Orsini N. Goodness of fit tools for dose-response meta-analysis of binary outcomes. *Res Synth Methods* 2017;8:149–60.
- 22 Schwartz J, Zanobetti A. Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death. *Epidemiology* 2000;11:666–72.
- 23 Hartemink N, Boshuizen HC, Nagelkerke NJD, *et al.* Combining risk estimates from observational studies with different exposure cutpoints: a meta-analysis on body mass index and diabetes type 2. *Am J Epidemiol* 2006;163:1042–52.