Author's response to reviews

Title: Synthesis of clinical prediction models under different sets of covariates

Authors:

Daisuke Yoneoka Mr (yoneoka.daisuke@ism.ac.jp)
Masayuki Henmi Mr (henmi@ism.ac.jp)
Norie Sawada Ms (nsawada@ncc.go.jp)
Manami Inoue Ms (mnminoue@m.u-tokyo.ac.jp)

Version: 5 Date: 5 August 2015

Author's response to reviews: see over
Manuscript title: Synthesis of clinical prediction models under different sets of covariates: a multivariate meta-analysis approach
Authors: Daisuke Yoneoka, Masayuki Henmi, Norie Sawada and Manami Inoue

Dear editors of *BMC Medical Research Methodology*

I am writing to submit a revision of the above-titled manuscript. The manuscript has been thoroughly and carefully revised to reflect the concerns outlined by the referees.

I am very grateful for the rigorous peer-review of our manuscript. I also highly appreciate the favorable comments and suggestions from the reviewers stating that this study is very relevant to important problems of current meta-analysis. I have attached a carefully structured response letter including point-by-point responses to the reviewer’s concerns.

I hope that the reviewers will find our response satisfactory, and look forward to hearing your judgment of the manuscript’s suitability for publication.

Yours sincerely,

Daisuke Yoneoka
School of Multidisciplinary Sciences
Department of Statistical Science,
The Graduate University for Advanced Studies,
10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
Email: yoneoka.daisuke@ism.ac.jp
Response letter

Manuscript ID: 8964180871673676
Manuscript title: "Synthesis of clinical prediction models under different sets of covariates: a multivariate meta-analysis approach"

We would like to thank the reviewers for their helpful comments. Their comments (italicized) are addressed on a point-by-point basis below.

Response to editorial corrections

1. We recommend that you copyedit the paper to improve the style of written English. If this is not possible, you may need to use a professional language editing service.

We thank you for this comment. Following the recommendation, we have used a professional editing service to improve our English.

In addition, the manuscript title may be misleading for readers who are interested in multivariate meta-analysis. Thus, we have changed the title from

“Synthesis of clinical prediction models under different sets of covariates: a multivariate meta-analysis approach”

to

“Synthesis of clinical prediction models under different sets of covariates.”
Response to reviewer 1

Major Compulsory Revisions

1. The main limitation of the paper is that the methods assume no heterogeneity between studies (either in the regression coefficients or in the joint distribution of covariates). This is very implausible in practice. While the authors do discuss the possibility of allowing for heterogeneity, I think it will be difficult to extend the methods to allow for heterogeneity in the regression coefficients and impossible (because of only 1 set of IPD) to allow for heterogeneity in the joint distribution of covariates. I think the authors should:
   a. discuss this limitation more
   b. include this limitation clearly in the abstract.
   c. moderate the conclusion "This study showed how to conduct the meta-analysis of regression coefficients under different covariate sets" (l309-310)

We thank the reviewer for these valuable suggestions for improving the quality of our manuscript. We agree that heterogeneity is an important consideration. To clarify this limitation, we have altered the following sentence from
Abstract line 24-26 (in the original manuscript):
   “These methods would be also useful for incorporating prior published information into the constructions of new prediction models.”

To
Abstract line 25-27 (in the revised manuscript):
   “If the assumption of homogeneity within studies is plausible, this methodology would be useful for incorporating prior published information into the construction of new prediction models.”

Other changes are listed below.
Line 150 (in the original manuscript):
   “and \( \hat{\theta}_i \) is the column vector of reported coefficients in the ith study.”
was altered to
Line 163-166 (in the revised manuscript):


“and \( \hat{\theta}_i \) is the column vector of reported coefficients in the \( i \)th study. The function \( f() \) comes from the omitted variable bias formula introduced in the previous section, whose formulation is reasonable if an assumption of homogeneity of studies in meta-analysis is acceptable.”

Line 188-189 (in the original manuscript):

“Here we assume a fixed effect model which presumes that there are no heterogeneity in the distribution of covariates and in values of parameters of interest.”

was altered to

Line 205-209 (in the revised manuscript):

“Here, we assume a fixed effect model which presumes that there is no heterogeneity in the distribution of covariates and in the values of the parameters of interest. This assumption may sometimes be unrealistic. Therefore, we recommend considering whether this assumption is reasonable based on background knowledge or reported information.”

Line 309-310 (in the original manuscript):

“This study showed how to conduct the meta-analysis of regression coefficients under different covariate sets”

was altered to

Line 350-354 (in the revised manuscript):

“This study demonstrated a method to conduct the meta-analysis of regression coefficients with different covariate sets under the assumption of homogeneity of studies (i.e., it is applicable in cases where studies in the meta-analysis have similar distributions of covariates and outcomes).”

2. Debray et al (2012) use a second missing data method – as well as mean imputation, they use "Uninformative regression coefficients", which I think just amounts to saying that missing regression coefficients are missing data and adapting the likelihood accordingly. Since mean imputation is obviously a bad idea, it would be good if the authors could also compare their method with "Uninformative regression coefficients" - although it is clear that "Uninformative regression coefficients" fails
to allow for omitted variable bias / confounding.

We thank the reviewer for this suggestion. Debray et al. (2012) imputed 0 to unknown coefficients with large variance (i.e., they assumed that the reported variance of coefficients is 100). They compared this approach, which they defined as “Uninformative regression coefficients,” with the mean imputation method. In fact, as recommended by the reviewer, we tried to include the results of this imputation method, but the results did not noticeably differ from M2 (mean imputation). Therefore, the M3 results (uninformative regression coefficients) were excluded from our report. The M3 results are listed in the table below. To clarify this point, we have revised our manuscript from

Line 23–25 (in the original manuscript):

“Debray et al. adopted an ad-hoc approach that uses mean imputation for the missing coefficient estimates to apply the multivariate meta-analysis method straightforwardly.”

to

Line 27-31 (in the revised manuscript):

“On the other hand, instead of using every IPD record, Debray et al. considered a method that uses the reported summary statistics from one set of IPD. They adopted an ad-hoc approach utilizing mean or zero imputations for the missing coefficient estimates to straightforwardly apply the multivariate meta-analysis method.”

Other changes are listed below:

Line 63-64 (in the original manuscript):

“Debray et al. utilize the mean imputation method for omitted coefficients and apply the technique of multivariate meta-analysis.”

was altered to

Line 69-71 (in the revised manuscript):

“To synthesize these regression coefficients, Debray et al. utilize the mean or zero imputation methods for omitted coefficients and apply the technique of multivariate meta-analysis.”
Line 229-230 (in the original manuscript):

“For example, coefficients and estimated standard errors of X2 from 3 studies (i=4,5,6) were imputed by the mean of other 6 studies.”

was altered to

Line 253-258 (in the revised manuscript):

“For example, coefficients and their estimated standard errors of X2 from 3 studies (i = 4, 5, 6) were imputed by the means of the other 6 studies. We tried the zero imputation method, which Debray et al. adopted and called uninformative regression coefficients, but it did not show notable results compared with the results from M2 (mean imputation). Therefore, we decided not to include the results of this method.”
Table: Performance of uninformative regression coefficients

<table>
<thead>
<tr>
<th>Case 1: Both X1 and X2 are continuous</th>
<th>Case 2: X1 is continuous and X2 is binary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Correlation r = 0</strong></td>
<td><strong>Correlation r = 0.5</strong></td>
</tr>
<tr>
<td>$-2$</td>
<td>$-1$</td>
</tr>
<tr>
<td><strong>M3: Uninformative regression coefficients</strong></td>
<td></td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>1.399</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-1.775</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.935</td>
</tr>
<tr>
<td><strong>M3: Uninformative regression coefficients</strong></td>
<td></td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>1.994</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>3.171</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.881</td>
</tr>
</tbody>
</table>
3. *In the simulation study, Table1, Case1, r=0, why is method M2 biased for alpha1 under the null? (bias reported as -0.171). In this setting, the expectation of alpha1 is 0 whether or not X2 is in the model, so surely the M2 method should also have expectation 0? I wonder if this suggests problems in the simulation.*

We greatly appreciate the valuable comment on the simulation study. This is a very important discussion point. The simulation was executed under the following three models:

(1) Logit \( P(Y = 1|X_1, X_2) = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 \)

(2) Logit \( P(Y = 1|X_1) = \beta_0 + \beta_1 X_1 \)

(3) Logit \( P(Y = 1|X_2) = \gamma_0 + \gamma_2 X_2 \).

Nine artificial studies were created and it is assumed for each study to report their estimated coefficients \((\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_0, \hat{\beta}_1, \hat{\gamma}_0, \hat{\gamma}_2)\) and their variances \((\text{Var}(\hat{\alpha}_0), \text{Var}(\hat{\alpha}_1), \text{Var}(\hat{\alpha}_2), \text{Var}(\hat{\beta}_0), \text{Var}(\hat{\beta}_1), \text{Var}(\hat{\gamma}_0), \text{Var}(\hat{\gamma}_2))\).

Under this assumption, the bias mentioned by the reviewer (in Case 1 with \( r = 0 \) and \( \alpha_1 = 0 \)) comes mainly from the biased estimator of \( \beta_0 \) compared with the true parameter \( \alpha_0 \). In this setting (Case 1, \( r = 0 \) and \( \alpha_1 = 0 \)), the estimates \( \hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\gamma}_0 \) and \( \hat{\gamma}_2 \) are not biased against the true parameters \((\alpha_0, \alpha_1, \alpha_2)\). The only biased estimate is \( \hat{\beta}_0 \)

(computed as \( \hat{\beta}_0 \approx \frac{\alpha_0 + 2}{\sqrt{1 + c^2}} \), where we set \( \alpha_2 = 1, \Omega_{2|X} = 1, c = \frac{16\sqrt{7}}{15\pi} \) in Eq. (4) of the manuscript). Since the M2 strategy is a multivariate meta-analysis after mean imputation, the biased estimate \( \hat{\beta}_0 \) influences the other synthesized coefficients (i.e., it influences the synthesized result \( \alpha_1 = -0.171 \)). In addition, the reported simulated variances of the coefficients are the diagonal elements in the multivariate meta-analysis. Since \( \hat{\beta}_0 \) is biased, \( \text{Var}(\hat{\beta}_0) \) is also biased. Therefore, the bias in \( \text{Var}(\hat{\beta}_0) \) also influences the synthesized result.

Although the true parameters cannot be obtained in practice, the true case can be checked in the simulation setting. In fact, if we correct \( \hat{\beta}_0 \) and \( \text{Var}(\hat{\beta}_0) \) and rerun the simulation, we obtain the results in the following table.
Table: Simulation results based on the true variance structure

<table>
<thead>
<tr>
<th>Case 1: Both X1 and X2 are continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation = 0</td>
</tr>
<tr>
<td>$\alpha_0 = 0$</td>
</tr>
<tr>
<td>Bias M2: Mean imputation</td>
</tr>
</tbody>
</table>
| $\alpha_0$                           | $-0.163$  
| $\alpha_1$                           | $-0.051$  
| $\alpha_2$                           | $0.061$  
| MSE M2: Mean imputation              |  
| $\alpha_0$                           | $0.138$  
| $\alpha_1$                           | $0.020$  
| $\alpha_2$                           | $0.051$  

According to this table, the correction reduces the bias of $\alpha_1$ from −0.171 to −0.051. To improve the quality of our manuscript, we have clarified this point by the following changes:

Line 234-235 (in the original manuscript):

"MSE of our estimator ranged from 0.021 to 0.803 (mean: 0.124) for Case 1 and from 0.012 to 0.486 (mean: 0.091) for Case 2."

Line 265-268 (in the revised manuscript):

"The MSE of our estimator ranged from 0.021 to 0.803 (mean: 0.124) for Case 1 and from 0.012 to 0.486 (mean: 0.091) for Case 2. Although the M2 strategy in Case 1 and $r = 0$ yielded somewhat biased results, the greatest amount of variation seemed to arise from the biased estimates of $\alpha_0$ in the models from which X2 was omitted."

Minor Essential Revisions

4. Ref [6] is central to the paper, but it's incomplete. I think the Stats in Medicine 2012 paper is meant (with a slightly different author set).

We thank the reviewer for this comment. We changed the reference from Line 406-407 (in the original manuscript):


to Line 457-459 (in the revised manuscript):

comparison of different approaches. Statistics in medicine 31(23), 2697–2712 (2012)"

5. There are a number of errors in the English: e.g. "the only" (abstract line3); "interests" (l12); "least squares" not "least square" (various places); "withs" (l49); and many more.

We thank the reviewer for pointing out the errors in the English. Based on this recommendation, we have checked the manuscript for such mistakes and employed a professional English editing service.

6. "Binalized" is not a word; "ordinal" is used where "ordinary" is meant; "Brier" should be used not "Breir".

We thank the reviewer for these corrections. As mentioned above, we have used a professional English editing service.

7. It is clear that the method proposed for continuous Z (page4) can be applied also for any Z, and I think this is what you do in the simulation study: please state this explicitly.

Thank you for the comment. The method proposed for continuous Z assumes that the conditional distribution of Z given X is a normal distribution. Accordingly, we can calculate the conditional expectation of Y given X (line 135-136 in the revised manuscript). Although Eq. (4) (in the manuscript) can be applied as a special case in our simulation, Eq. (3) (in the manuscript) is more general and is applicable to any distribution of covariates. Therefore, in the simulation, we empirically calculate the omitted variable bias by Eq. (3) rather than by Eq. (4). To improve the quality of our manuscript, we have clarified this point by making the following changes:
Line 196-197 (in the original manuscript):

“In this section, a Monte Carlo simulation was performed to evaluate the performances of our proposed method.”

to
Line 219-221 (in the revised manuscript):
“"In this section we describe a Monte Carlo simulation which was performed to evaluate the performance of our proposed method. In the simulation, we empirically calculate the omitted variable bias by using equation (3) instead of equation (4).”

8. Similarly, the method is described (p5) for just one pattern of incomplete covariates: please state explicitly that the method generalises easily to many patterns.

We thank the reviewer for this valuable suggestion. We agree that the generality of the method needs to be explicitly stated. To clarify this point, we have revised the following sentence from Line 140-141 (in the original manuscript):
“Here we show only the case where Z is omitted, but the X omitted case can be considered in the same manner.”

to Line 152-154 (in the revised manuscript):
“Here we show only the case where Z is omitted, but the case where X is omitted can be considered in the same manner, and further, it is easy to generalize to various other omittance patterns.”

9. The asymptotics considered on page 6 are not useful since they refer to the number of studies going to infinity, but the number of studies does not get large in practice.

We appreciate this important discussion point. Although the number of studies rarely approaches infinity in practice, we consider that the discussion on the asymptotic property is necessary to validate our method from a theoretical perspective. In addition, one of the purposes of the simulation was to check this asymptotic property in an approximate setting. Therefore, we have decided to maintain the asymptotic analysis in our revised manuscript.

10. In the simulation design, with bivariate uncorrelated X1, X2 and alpha1=1, I estimate p(Y)=0.983. Given the sample size of 100, this suggests only 2 events on
average. What happens in the simulation study when the models cannot be fitted?

Thank you for this valuable comment. In the simulation, we excluded extreme values because they yielded extreme $P(Y)$ at $\alpha_1 = 1$, and thus sometimes failed to converge. Consequently, such extreme values were excluded when calculating the bias and mean squared errors (MSE). To clarify this point, we have revised our manuscript from

Line 223-224 (in the original manuscript):

“Under these settings, 1000 Monte Carlo simulations were implemented.”

to

Line 245-246 (in the revised manuscript):

“Under these settings, 1000 Monte Carlo simulations were implemented. If the models could not be fitted and converged, their results were excluded from the calculation of bias and mean squared error (MSE).”

II. How were "sporadically missing" data handled - i.e. studies with covariates that were partly observed?

This is a pertinent question. In the application analysis, we needed to handle “sporadically missing” data in several areas. For this purpose, we implemented a complete case analysis, in which samples with partly missing covariates were excluded from the final dataset. The average percentage of missing observations was 2.8% (median: 1.5%; SD: 2.5%), indicating that omitting these data had little effect on the results. To clarify this point, we have revised the text from

Line 278-279 (in the original manuscript):

“Coefficients from each model were stored as the aggregated statistics, which could be regarded as prior studies for meta-analysis.”

to

Line 313-316 (in the revised manuscript):

“Coefficients from each model were stored as aggregated statistics, which could be regarded as prior studies for meta-analysis. In terms of handling sporadically missing data (average missing rate was 2.8% with a standard deviation of 2.5%), complete case analysis was executed.”
12. Table 2, have AUC and Brier score been multiplied by 100?

Thank you for this comment. These metrics were indeed multiplied by 100. In addition, as recommended by reviewer 2, we added a calibration accuracy measure (the Hosmer–Lemeshow chi-squared statistic). Consequently, we have altered the sentence

Line 285-288 (in the original manuscript):

“The performance was measured by the area under the receiver operator characteristic curve (AUC) and the Brier score (BS), which are one of indicators of accuracy of the prediction model. Higher AUC means higher prediction accuracy and BS is vice versa.”

to

Line 325-328 (in the revised manuscript):

“The discriminant performance of the prediction models was measured by the area under the receiver operator characteristic curve (AUC) and the Brier score (BS) (multiplied by 100), both of which are indicators of the accuracy of the prediction model. A higher AUC indicates higher prediction accuracy, while the BS has an inverse relationship. In addition, the model’s calibration was examined by the Hosmer-Lemeshow chi-squared statistic.”

Discretionary Revisions

13. Some other methods have been proposed for addressing similar problems and could usefully be mentioned:


Thank you for this recommendation. Since these papers are important and were also
mentioned by reviewer 2, we have included them to enhance the quality of our paper. Accordingly, we have altered the sentence

Line 23-25 (in the original manuscript):
   “Debray et al. adopted an ad-hoc approach that uses mean imputation for the missing coefficient estimates to apply the multivariate meta-analysis method straightforwardly.”

and

Line 23-31 (in the revised manuscript):
   “To tackle this problem, the Fibrinogen Studies Collaboration proposed a multivariate meta-analysis approach to borrow strength from partially adjusted results by using individual patient data (IPD), and Riley et al. demonstrated the approach in practice. In addition, Resche-Rigon et al. adopted a multiple imputation method with IPD. On the other hand, instead of using every IPD record, Debray et al. considered a method that uses the reported summary statistics with one set of IPD. They adopted an ad-hoc approach utilizing mean or zero imputations for the missing coefficient estimates to straightforwardly apply the multivariate meta-analysis method.”

14. The problem is described as “omitted variable bias”. As a result I wondered until page 4 whether it was going to assume the omitted variables were uncorrelated with the included variables. I would clarify this by stressing early on that the method addresses confounding.

Thank you for the suggestion. We have accordingly altered the following sentence from

Line 71-72 (in the original manuscript):
   “which means the prior models have the subset of covariates and are considered as "under-specified" models.”

and

Line 78-81 (in the revised manuscript):
   “which means the prior models have subsets of covariates and are considered as under-specified models. Note that since the omitted variables from the true models (full models) are correlated with the included variables, the subset
models are confounded and biased compared with true models.”

15. You use Greek symbols for parameters, except for \( p_{XZ} \): would you consider using Greek here too?

Thank you for the recommendation. In this study, \( p_{XZ} \) represents whole parameters related to the joint distribution of covariates. If the joint distribution is multivariate normal, we can easily replace \( p_{XZ} \) with mean and variance parameters. However, it is generally difficult to list all parameters concerning the distribution of covariates. Therefore, we adopt the \( p_{XZ} \) notation as a general expression of the distribution.
Response to reviewer 2

1. My major comment is that this does not extend the existing literature. Where systematically missing confounders (adjustment factors) are not available in some studies, a multivariate meta-analysis approach has already been proposed to borrow strength from partially adjusted results. Please see [1] where this is proposed and also [2] where this is demonstrated. I therefore do not see what this paper is added, apart from looking at the change in C-statistic and Brier score in the example.

We greatly appreciate the reviewer’s suggestions for improving the quality of our manuscript. These points are worthy of discussion. Although [1] and [2] are important papers and closely related to our study, there are several differences between our method and their method, as described below:

1) The method in [1] and [2] assumes that all individual patient data (IPD) are available. Our method uses mainly summary statistics (reported regression coefficients and their variance) with single IPD.

2) Paper [1] applies both full and partial models to the IPD of each cohort (if the full covariates are available). For a given cohort, the coefficients in the full and partial models ($\beta_1$, $\beta_2$ and $\beta_1^p$, respectively) were estimated from the same IPD. Therefore, the correlation (denoted as $\rho$) between $\hat{\beta}_1^f$ and $\hat{\beta}_1^p$ is naturally introduced ($\beta_1^p$ is only estimated in the cohort with subsets of covariates). On the other hand, our approach assumes that each cohort independently reports either $\hat{\beta}_1^f$, $\hat{\beta}_2$ or $\hat{\beta}_1^p$; thus, the correlation parameter $\rho$ has no role in our method. Because we assume homogeneity among studies, the misspecified models are related to the true models. (i.e., Coefficients in the misspecified models include information of coefficients in the true models.) Our approach utilizes this relationship to improve the efficiency of estimating the true parameters.

To improve the quality of our manuscript, we have included these important papers ([1] and [2]) and altered the manuscript from Line 23-25 (in the original manuscript):

“Debray et al. adopted an ad-hoc approach that uses mean imputation for the
missing coefficient estimates to apply the multivariate meta-analysis method straightforwardly.”

to

Line 23-31 (in the revised manuscript):

“To tackle this problem, the Fibrinogen Studies Collaboration proposed a multivariate meta-analysis approach to borrow strength from partially adjusted results by using individual patient data (IPD), and Riley et al. demonstrated the approach in practice. In addition, Resche- Rigon et al. adopted a multiple imputation method with IPD. On the other hand, instead of using every IPD record, Debray et al. considered a method that uses the reported summary statistics with one set of IPD. They adopted an ad-hoc approach utilizing mean or zero imputations for the missing coefficient estimates to straightforwardly apply the multivariate meta-analysis method.”

Further, we changed the manuscript from

Line 317-318 (in the original manuscript):

“(but such prior results are just reported in the form of summary statistics).”

to

Line 361-366 (in the revised manuscript):

“(but with such prior results reported just in the form of summary statistics). The minimal use of IPD (use of one IPD and other summary statistics) distinguishes our approach from that of the Fibrinogen Studies Collaboration. They assume that both full and partial models are applied in each cohort by using its cohort IPD, and thus the estimation of the correlation of coefficients between full and partial models is applicable.”

2. Also, I did not find the paper easy to follow. Though the question is well explained early on, I cannot easily see if the methods how the multivariate approach is accounting for studies that do not present all covariates of interest.

Thank you for these valuable comments. We used a professional editorial service to improve the clarity of our manuscript.

In our method, the omitted variable biases are functions of the true parameters in the
full models, and the bias formula imparts strength from other subset models.
Therefore, we can use the indirect information from studies that do not present all covariates.
Since our approach synthesizes multiple outcomes (multiple coefficients), we adopted the term “multivariate meta-analysis” in our manuscript. However, as mentioned by the reviewer, our method differs from standard multivariate meta-analysis. To avoid confusing readers in this field, we have altered the title from “Synthesis of clinical prediction models under different sets of covariates: a multivariate meta-analysis approach” to “Synthesis of clinical prediction models under different sets of covariates”

3. I also do not follow the rationale for the non-linear modelling of the covariates. Why not just assume normality of the original scale of the alpha and beta coefficients? This is what the authors did in [1], as the regression coefficients are MLE estimates and so should be asymptotically normal I think?

We thank the reviewer for this comment. Our method directly estimates the parameters in the true model (for simplicity, we set the true model as the full model and the other models as subsets of this true model). Because we assume homogeneity among studies, the parameters in the subset models are related to those of the full models (true models). To represent the parameter relationships between the full and subset models, we require non-linear modeling of the covariates because of the logistic regression formulation. To clarify this point, we have altered the sentence Line 71-72 (in the original manuscript):

“which means the prior models have the subset of covariates and are considered as ”under-specified” models.”

to

Line 78-81 (in the revised manuscript):

“which means the prior models have subsets of covariates and are considered as under-specified models. Note that since the omitted variables from the true models (full models) are correlated with the included variables, the subset
models are confounded and biased compared with true models.”

We also changed
Line 150 (in the original manuscript):
“and $\hat{\theta}_i$ is the column vector of reported coefficients in the ith study.”
to
Line 163-166 (in the revised manuscript):
“and $\hat{\theta}_i$ is the column vector of reported coefficients in the ith study. The function $f()$ comes from the omitted variable bias formula introduced in the previous section, whose formulation is reasonable if an assumption of homogeneity of studies in meta-analysis is acceptable.”

4. Critically, the multivariate approach presented does not appear to be a random effects model. I doubt whether the fixed effect approach is realistic.

The reviewer 1 has pointed out one limitation of our approach. Accordingly, we have revised our manuscript to clarify this limitation and moderated our conclusion as follows:
Abstract line 24-26 (in the original manuscript):
“These methods would be also useful for incorporating prior published information into the constructions of new prediction models.”
was altered to
Abstract line 25-27 (in the revised manuscript):
“If the assumption of homogeneity within studies is plausible, this methodology would be useful for incorporating prior published information into the construction of new prediction models.”

Line 188-189 (in the original manuscript):
“Here we assume a fixed effect model which presumes that there are no heterogeneity in the distribution of covariates and in values of parameters of interest.”
was altered to
Line 205-209 (in the revised manuscript):
“Here, we assume a fixed effect model which presumes that there is no heterogeneity in the distribution of covariates or in the values of the parameters of interest. This assumption may sometimes be unrealistic. Therefore, we recommend considering whether this assumption is reasonable based on background knowledge and reported information.”

Line 309-310 (in the original manuscript):

“This study showed how to conduct the meta-analysis of regression coefficients under different covariate sets”

was altered to

Line 350-354 (in the revised manuscript):

“This study demonstrated a method to conduct the meta-analysis of regression coefficients with different covariate sets under the assumption of homogeneity of studies (i.e., it is applicable in cases where studies in the meta-analysis have similar distributions of covariates and outcomes).”

5. The authors repeatedly say that the method that ignores the studies without the full set of covariates is biased or wrong. Why? If these are a random sample, then why would they not give unbiased regression coefficients and thus the right prediction model (on average)?

Thank you for this valid comment. In our manuscript, we do not intend to insist that methods that ignore the studies without the full set of covariates are biased or wrong. We intended to describe that including studies with common covariate sets would reduce the efficiency of the multivariate meta-analysis, because it would ignore the indirect information from the models with different covariate sets. In fact, in the simulation, we checked that our method (called “M1: Full set only” in the simulation section) is asymptotically non-biased. Including the indirect information is one purpose of our study. Debray et al. applied the multivariate meta-analysis approach using mean or zero imputation methods, which lead to bias. Therefore, we propose correcting this bias with an omitted bias formula, which improves the efficiency of the analysis. To clarify this point, we have made the following adjustments:

Line 66-67 (in the original manuscript):
“However, the former approach leads to biased results and the latter ignores the indirect information from omitted studies.”

was altered to

Line 73-75 (in the revised manuscript):

“However, the former approach leads to biased results and the latter is not biased but leads to loss of efficiency by ignoring indirect information from omitted studies.”

Line 226-227 (in the original manuscript):

“M1 was multivariate meta-analysis using only 3 studies with a full set of covariates.”

was altered to

Line 248-251 (in the revised manuscript):

“M1 was the multivariate meta-analysis using only 3 studies with a full set of covariates. From a theoretical perspective, the M1 strategy does not include any bias but is inferior in efficiency compared with our proposed method, which can be checked using the results of MSE.”

6. The sandwich estimator approach is interesting for dealing with mis-specified within-study correlations, but again how does this build on the existing literature such as [3]?

Thank you for this comment. We missed Reference [3], which was only recently published. Our proposed sandwich estimator is same with the method of Liu et al. [9] and is very relevant to the method proposed in [3]. However, although all three methods yield very similar results, our method (and Liu et al. [9]) differs from that of [3] in being a two-stage procedure; 1) it estimates the synthesized coefficients, placing zeros in the off-diagonal elements of the covariance matrix (i.e., it assumes zero correlation between the estimated coefficients) and 2) it inputs the first-stage estimates to the sandwich covariance matrix.

This small difference exists in the implementation; conceptually, our method and Liu et al.’s method [9] are very similar to that proposed in [3]. To clarify this point, we have altered
Line 345-347 (in the original manuscript):

“Diag[x] is the diagonal matrix with the elements of x on the diagonal.”

to

Line 196-197 (in the revised manuscript):

“Diag[x] is the diagonal matrix with the elements of x on the diagonal. This idea essentially comes from Liu et al. and can be regarded as analogy of the result proposed by Chen et al.”

7. The simulation studies show minimal improvement in the performance of the prediction model. This is expected, given that I think the regression coefficients should be unbiased for all approaches. The authors note this in their discussion, but say that there is considerable gain in precision. This is interesting – but again a finding that is already well known in this field, shown in many papers that the authors do not seem to be aware of, such as [4-6] (sorry to be referencing a number of my own papers here, but they do seem very relevant)

We appreciate your comments and agree that the gain in precision is a well-known result. The papers recommended by the reviewer are indeed relevant to our study. To clarify this issue, we have altered the following sentence

Line 345-347 (in the original manuscript):

“Although the improvement of accuracy of the prediction model was relatively small, confidence intervals of synthesized coefficients were dramatically decreased because information from other studies helped to gain an efficiency.”

to

Line 395-399 (in the revised manuscript):

“Although the improvement of accuracy of the prediction model was relatively small, the confidence intervals of synthesized coefficients were dramatically decreased because information from other studies helped improve efficiency. In the context of multivariate meta-analysis, it is well known that we can gain precision by borrowing strength from other partially reported results.”
8. Finally, the prediction model’s performance considers discrimination but not calibration explicitly (e.g., calibration slope or calibration in the large), so this is not a complete evaluation of the model performance statistically [7].

Thank you for this suggestion. We absolutely agree on the importance of model calibration. A common measure of calibration is the Hosmer–Lemeshow’s chi-squared statistic, which compares the observed and predicted outcomes over deciles of risk [7, 8]. We have included the calibration measure in our revised manuscript. Consequently

Line 285-288 (in the original manuscript):

“The performance was measured by the area under the receiver operator characteristic curve (AUC) and the Brier score (BS), which are one of indicators of accuracy of the prediction model. Higher AUC means higher prediction accuracy and BS is vice versa.”

was altered to

Line 323-328 (in the revised manuscript):

“The discriminant performance of the prediction models was measured by the area under the receiver operator characteristic curve (AUC) and the Brier score (BS) (multiplied by 100), both of which are indicators of the accuracy of the prediction model. A higher AUC indicates higher prediction accuracy, while the BS has an inverse relationship. In addition, the model’s calibration was examined by the Hosmer-Lemeshow chi-squared statistic.”

Moreover, the Hosmer–Lemeshow statistic was incorporated into Table 2 (page 15 of the revised manuscript). Part of the revised table is shown below:

Table: Estimated regression coefficients (and standard errors) from JPHC data (partly shown)

<table>
<thead>
<tr>
<th></th>
<th>Area 1</th>
<th>Area 2</th>
<th>Area 3</th>
<th>Area 4</th>
<th>Area 5</th>
<th>Area 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>67.01</td>
<td>68.74</td>
<td>67.97</td>
<td>65.52</td>
<td>69.16</td>
<td>68.19</td>
</tr>
<tr>
<td>Brier score</td>
<td>7.71</td>
<td>7.72</td>
<td>7.65</td>
<td>8.07</td>
<td>7.78</td>
<td>7.68</td>
</tr>
<tr>
<td>Hosmer–Lemeshow</td>
<td>10.91</td>
<td>15.78*</td>
<td>14.01</td>
<td>101.45*</td>
<td>56.96*</td>
<td>17.70*</td>
</tr>
<tr>
<td></td>
<td>Area 7</td>
<td>Area 8</td>
<td>Area 9</td>
<td>Area 10</td>
<td>Proposed</td>
<td>Conventional</td>
</tr>
<tr>
<td>---------------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>---------</td>
<td>----------</td>
<td>--------------</td>
</tr>
<tr>
<td>AUC</td>
<td>68.32</td>
<td>69.47</td>
<td>68.29</td>
<td>67.28</td>
<td>68.01</td>
<td>67.24</td>
</tr>
<tr>
<td>Brier score</td>
<td>7.77</td>
<td>7.63</td>
<td>7.64</td>
<td>8.03</td>
<td>7.72</td>
<td>7.8</td>
</tr>
<tr>
<td>Hosmer–Lemeshow</td>
<td>58.60*</td>
<td>28.70*</td>
<td>25.38*</td>
<td>186.76*</td>
<td>21.13*</td>
<td>21.17*</td>
</tr>
</tbody>
</table>

* p-value of Hosmer–Lemeshow test is less than 0.05.

References


