

RESEARCH ARTICLE OPEN ACCESS

Modeling the Role of Baseline Risk and Additional Study-Level Covariates in Meta-Analysis of Treatment Effects

Phuc T. Tran  | Annamaria Guolo 

Department of Statistical Sciences, University of Padova, Padova, Italy

Correspondence: Phuc T. Tran (thienphuc.tran@studenti.unipd.it)**Received:** 24 April 2024 | **Revised:** 8 June 2025 | **Accepted:** 13 September 2025**Funding:** The authors received no specific funding for this work.**Keywords:** baseline risk | heterogeneity | likelihood | measurement error | meta-analysis

ABSTRACT

The relationship between the treatment effect and the baseline risk is a recognized tool to investigate the heterogeneity of treatment effects in meta-analyses of clinical trials. Since the baseline risk is difficult to measure, a proxy is adopted, which is based on the rate of events for the subject under the control condition. The use of the proxy in terms of aggregated information at the study level implies that the data are affected by measurement errors, a problem that the literature has explored and addressed in recent years. This paper proposes an extension of the classical meta-analysis with baseline risk information, which includes additional study-specific covariates other than the rate of events to explain heterogeneity. Likelihood-based inference is carried out by including measurement error correction techniques necessary to prevent unreliable inference due to the measurement errors affecting the covariates summarized at the study level. Within-study covariances between risk measures and the covariate components are computed using Taylor expansions based on study-level covariate subgroup summary information. When such information is not available and, more generally, in order to reduce computational difficulties, a pseudo-likelihood solution is developed under a working independence assumption between the observed error-prone measures. The performance of the methods is investigated in a series of simulation studies under different specifications for the sample size, the between-study heterogeneity, and the underlying risk distribution. They are applied to a meta-analysis about the association between COVID-19 and schizophrenia.

1 | Introduction

One common application of meta-analysis is the evaluation of the effectiveness of a treatment, exploiting information from different studies comparing a treatment group and a control group [1–3]. A proper analysis needs to detect and account for between-study heterogeneity, which occurs as a consequence of different sources, including differences in studies' protocols and design, execution, and experience of the staff, as well as differences in patients' characteristics and general health, such as

age, gender, and severity of illness. If individual patient data are available from each study included in the meta-analysis, then an accurate investigation of the relationship between treatment risk and features of patients and/or studies is possible. Unfortunately, information is usually available in terms of aggregated data at the study level, that is, in terms of summary results from published studies [4, 5]. Pioneering studies in the 1990s suggest that a relevant source of heterogeneity is represented by the underlying risk or the baseline risk of disease of study participants; see, for example, Sharp [2] and Schmid et al. [3]. This kind of information

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Statistics in Medicine* published by John Wiley & Sons Ltd.

is not available at the patient level, and it is also difficult to measure, since it can be seen as an aggregation of relevant attributes of the patient and the severity of illness. Accordingly, the baseline risk is approximated by using a surrogate measure on patients under the control condition, that is, a measure of the rate at which the outcome of interest occurs. The inclusion of a study-specific measure of the baseline risk for the controls in the meta-analysis model in order to explain between-study heterogeneity gives rise to a meta-regression model [6–8]. Since the risk measured on controls is aggregated study-level information used as a proxy for the true unknown baseline risk, it is affected by error [1, 7, 9, 10]. The presence of measurement error can affect inference and give rise to inconsistent results, as the first studies in McIntosh [1] and Sharp and Thompson [11] show. See also the discussion in Arends et al. [6], van Houwelingen et al. [7] and Guolo et al. [9]. More generally, outside the meta-analysis framework, a substantial and long-established literature has shown that the consequences of measurement errors in linear and nonlinear regression models can be substantial if errors are not corrected for. See the book-length discussion in Carroll et al. [12], Yi [13], and Yi et al. [14]. The best known effect is attenuation bias, that is, a biased toward zero estimate of the coefficient associated with the risk measure in the control group, which occurs under an additive and homoscedastic between-study error on the underlying risk measure; see, for example, van Houwelingen et al. [7] and Guolo [10]. Investigations in meta-analysis, including a measure of the baseline risk, have provided several solutions to correct for the presence of measurement errors. Methods include structural and functional approaches, according to the choice between accepting assumptions on the underlying unobserved control measure (structural approaches) or not (functional approaches). See Ghidey et al. [15] and Guolo [10] for a comparison of some of the techniques, including likelihood-based solutions under different assumptions of the underlying risk, corrected and conditional scores, and a simulation-based approach [16].

Traditionally, control risk as a measure of the baseline risk is the only covariate considered in the meta-regression model of treatment effectiveness to describe between-study heterogeneity. This paper proposes an extension that includes additional study-specific covariates other than the control risk, which might help to improve the explanation of heterogeneity in the treatment risk. Such additional covariates are summary quantities at the study level, for example, the mean age of patients or the percentage of females/males. As it happens, for the measure of the baseline risk, the available summarized covariates are a surrogate for the actual true measurement that would occur at the individual subject level in the study, and thus, they are likely affected by measurement errors. Appropriate inference should take this additional source of error, as well as the one associated with the baseline risk measure, into account. The paper focuses on a structural approach using likelihood-based correction techniques applied to a hierarchical model relating the unobserved error-prone quantities to their observed counterparts under an approximate or an exact measurement error specification. When the approximate normal error structure is chosen, mainly for computational convenience, within-study covariances between risk measures and the covariate components are computed by using Taylor expansions based on study-level covariate subgroup summary information. When such information is not available and, more generally, in order to reduce computational difficulties,

a pseudo-likelihood solution is developed under a working independence assumption between the observed error-prone measures. Under the exact binomial or multinomial measurement error structure, the likelihood function has no closed form, and it can be computationally intensive to maximize. The pseudo-likelihood version is considered as a viable option.

The performance of the proposed solutions is investigated in a series of simulation studies, including different types of additional covariates, either a mean response or a binary indicator of a feature from each study included in the meta-analysis. Scenarios cover different specifications for sample size, between-study heterogeneity, and underlying risk distribution. The applicability of the methods is also evaluated on a meta-analysis of the association between mortality from COVID-19 and schizophrenia.

2 | Classical Meta-Analysis Model With Baseline Risk Information

Consider a meta-analysis of n independent studies comparing a treatment group and a control group in order to investigate the effectiveness of a treatment. Let η_i and ξ_i denote the measure of risk in the treatment group and in the control group for study i , respectively, as, for example, log-odds or log-event rate. The common meta-analysis model, including baseline risk information, assumes that the quantities are related according to a linear model [6, 9],

$$\eta_i = \beta_0 + \beta_1 \xi_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \tau^2), \quad (1)$$

where τ^2 is the residual variance describing the variation among studies in the treatment risk η_i unexplained by the underlying risk ξ_i , for example, heterogeneity due to patient characteristics or study design. The sampling errors ε_i are independent of the underlying risk ξ_i . The inferential interest is usually in β_1 , which represents the relationship between η_i and ξ_i . The use of the control risk as a proxy for the unknown and unavailable baseline risk requires a model to be specified to relate the true/unobserved quantities and the mismeasured counterparts. Since the information from study i is an average value thought to be representative of the subjects included in that study, the measurement error model is specified through a classical structure, according to the measurement error terminology [12, 14]. This means that the error is additive and homoscedastic, with mismeasured values thought to vary randomly around the unobserved true value. The specification of the exact measurement error structure is strictly related to the available information from each study included in the meta-analysis. To fix the ideas, and without loss of generality, let Y_i and X_i denote the number of events in the treatment group and in the control group from study i , respectively, and let n_{iT} and n_{iC} denote the number of treated and the number of controls, respectively. Using the log-odds as a risk measure, the observed error-prone version of $(\eta_i, \xi_i)^T$ is given by

$$\hat{\eta}_i = \log\left(\frac{Y_i}{n_{iT} - Y_i}\right), \quad \hat{\xi}_i = \log\left(\frac{X_i}{n_{iC} - X_i}\right) \quad (2)$$

with associated standard error

$$s_{\hat{\eta}_i}^2 = \frac{1}{Y_i} + \frac{1}{n_{iT} - Y_i}, \quad s_{\hat{\xi}_i}^2 = \frac{1}{X_i} + \frac{1}{n_{iC} - X_i} \quad (3)$$

According to an approximate specification, the relationship between $(\hat{\eta}_i, \hat{\xi}_i)^T$ and $(\eta_i, \xi_i)^T$ is directly modeled, usually through a normal specification for computational convenience [1, 6],

$$\begin{pmatrix} \hat{\eta}_i \\ \hat{\xi}_i \end{pmatrix} | \begin{pmatrix} \eta_i \\ \xi_i \end{pmatrix} \sim N \left(\begin{pmatrix} \eta_i \\ \xi_i \end{pmatrix}, \begin{pmatrix} s_{\eta_i}^2 & 0 \\ 0 & s_{\xi_i}^2 \end{pmatrix} \right).$$

The exact measurement error model relating $(Y_i, X_i)^T$ to $(\eta_i, \xi_i)^T$ can be specified by considering the binomial distribution for Y_i and X_i [4, 6], namely,

$$\begin{aligned} Y_i | \eta_i &\sim \text{Binomial}\{n_{iT}, \text{expit}(\eta_i)\}, \\ X_i | \xi_i &\sim \text{Binomial}\{n_{iC}, \text{expit}(\xi_i)\}. \end{aligned}$$

See, for example, Guolo [9] for details and different specifications, such as, for example, the Poisson case.

Following the terminology of the measurement error literature, in this paper, we will consider a structural approach to deal with error-prone measures of $(\eta_i, \xi_i)^T$. Accordingly, the unobserved mismeasured quantities are interpreted as random variables, and a distribution function for them needs to be specified. Assume that the underlying risk ξ_i follows a Normal distribution, $\xi_i \sim N(\mu_\xi, \sigma_\xi^2)$, mainly for computational convenience. Accordingly, under the approximate measurement error model, the marginal distribution for $(\hat{\eta}_i, \hat{\xi}_i)^T$ is

$$\begin{pmatrix} \hat{\eta}_i \\ \hat{\xi}_i \end{pmatrix} \sim N_2 \left\{ \begin{pmatrix} \beta_0 + \beta_1 \mu_\xi \\ \mu_\xi \end{pmatrix}, \begin{pmatrix} s_{\eta_i}^2 + \beta_1^2 \sigma_\xi^2 + \tau^2 & \beta_1 \sigma_\xi^2 \\ \beta_1 \sigma_\xi^2 & s_{\xi_i}^2 + \sigma_\xi^2 \end{pmatrix} \right\},$$

and the likelihood function for the whole parameter vector $\theta = (\beta_0, \beta_1, \mu_\xi, \tau^2, \sigma_\xi^2)^T$ has a closed-form expression,

$$\begin{aligned} L(\theta) \propto \prod_{i=1}^n &\left\{ \det \begin{pmatrix} s_{\eta_i}^2 + \beta_1^2 \sigma_\xi^2 + \tau^2 & \beta_1 \sigma_\xi^2 \\ \beta_1 \sigma_\xi^2 & s_{\xi_i}^2 + \sigma_\xi^2 \end{pmatrix} \right\}^{-\frac{1}{2}} \\ &\phi_2 \left[\begin{pmatrix} s_{\eta_i}^2 + \beta_1^2 \sigma_\xi^2 + \tau^2 & \beta_1 \sigma_\xi^2 \\ \beta_1 \sigma_\xi^2 & s_{\xi_i}^2 + \sigma_\xi^2 \end{pmatrix}^{-\frac{1}{2}} \left\{ \begin{pmatrix} \hat{\eta}_i \\ \hat{\xi}_i \end{pmatrix} - \begin{pmatrix} \beta_0 + \beta_1 \mu_\xi \\ \mu_\xi \end{pmatrix} \right\} \right], \end{aligned}$$

where ϕ_2 denotes the density function of the bivariate standard normal distribution.

In case the exact measurement error model is specified, the likelihood function associated with the number of events $(Y_i, X_i)^T$ is not in closed-form,

$$\begin{aligned} L(\theta) \propto \prod_{i=1}^n &\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \{\text{expit}(\eta_i)\}^{Y_i} \{1 - \text{expit}(\eta_i)\}^{n_{iT}-Y_i} \\ &\{\text{expit}(\xi_i)\}^{X_i} \{1 - \text{expit}(\xi_i)\}^{n_{iC}-X_i} \\ &\times \frac{1}{\sqrt{\tau^2}} \phi \left(\frac{\eta_i - \beta_0 - \beta_1 \xi_i}{\tau} \right) \times \frac{1}{\sqrt{\sigma_\xi^2}} \phi \left(\frac{\xi_i - \mu_\xi}{\sigma_\xi} \right) d\eta_i d\xi_i, \end{aligned} \quad (4)$$

where ϕ denotes the density function of the standard normal distribution. The computation of the likelihood function can be done by using numerical methods, for example, Gauss–Hermite quadrature. Inference based on the exact measurement error model can be prone to some issues. Several authors warn against the risk of convergence problems, given by a nonpositive definite covariance matrix and values of the variance components on the boundary of the parameter space, especially in the case of small sample size [17–19].

3 | Including Additional Covariates

This paper investigates the extension of the classical meta-analysis model with baseline risk information to include study-specific covariates as a way to explain potential residual heterogeneity. The focus will be on a study-specific covariate ζ_i , which may or may not be affected by measurement error. This will give rise to different modifications of the meta-regression model and its measurement error component and, as a consequence, to modifications of the likelihood expression.

3.1 | Error-Free Covariates

Suppose that each study included in the meta-analysis provides information about features of the study or patients characteristics that are not affected by errors. Typically, such information is not a summarized quantity over the subjects included in the study; instead, it is a common quantity, for example, the country where the study has been conducted. In this case, ζ_i only enters the model with baseline risk information (1),

$$\eta_i = \beta_0 + \beta_1 \xi_i + \beta_2 \zeta_i + \varepsilon_i, \quad \xi_i \sim N(0, \sigma_\xi^2), \quad \varepsilon_i \sim N(0, \tau^2), \quad (5)$$

while the measurement error structure is not affected. Correspondingly, there is no additional computational effort in the likelihood-based analysis.

3.2 | Error-Affected Covariates

More often, the additional covariate represents aggregated information over the subjects included in each study, for example, the mean age or the percentage of females/males. The study-specific information $\hat{\zeta}_i$ is a summary of information of the true unknown covariate ζ_i , and thus it is affected by errors. The measurement error model relating $(\hat{\eta}_i, \hat{\xi}_i)^T$ to $(\eta_i, \xi_i)^T$ must be extended to include a distribution for $\hat{\zeta}_i$ given the true unknown quantity. Following the structural approach for measurement error correction, a marginal distribution for ζ_i must be specified. Choosing a Normal distribution $\zeta_i \sim N(\mu_\zeta, \sigma_\zeta^2)$, for computational convenience, the model for risk measures and the error-affected covariate can be expressed as

$$\eta_i = \beta_0 + \beta_1 \xi_i + \beta_2 \zeta_i + \varepsilon_i, \quad (6)$$

with

$$\xi_i \sim N(\mu_\xi, \sigma_\xi^2), \quad \zeta_i \sim N(\mu_\zeta, \sigma_\zeta^2), \quad \varepsilon_i \sim N(0, \tau^2).$$

For the moment, suppose we work with the approximate error model (2). Inserting the information from $\hat{\zeta}_i$ in the approximate error model leads to

$$\begin{pmatrix} \hat{\eta}_i \\ \hat{\xi}_i \\ \hat{\zeta}_i \end{pmatrix} \Big| \begin{pmatrix} \eta_i \\ \xi_i \\ \zeta_i \end{pmatrix} \sim N_3 \left(\begin{pmatrix} \eta_i \\ \xi_i \\ \zeta_i \end{pmatrix}, \Gamma_i = \begin{pmatrix} s_{\eta_i}^2 & 0 & s_{\eta_i, \zeta_i} \\ 0 & s_{\xi_i}^2 & s_{\xi_i, \zeta_i} \\ s_{\eta_i, \zeta_i} & s_{\xi_i, \zeta_i} & s_{\zeta_i}^2 \end{pmatrix} \right), \quad (7)$$

where $s_{\zeta_i}^2$ denotes the within-study variance of $\hat{\zeta}_i$, s_{η_i, ζ_i} denotes the within-study covariance between $\hat{\eta}_i$ and $\hat{\zeta}_i$, and s_{ξ_i, ζ_i} denotes the within-study covariance between $\hat{\xi}_i$ and $\hat{\zeta}_i$. The marginal distribution for the observed measures of risk and covariate is

$$\begin{pmatrix} \hat{\eta}_i \\ \hat{\xi}_i \\ \hat{\zeta}_i \end{pmatrix} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma} + \Gamma_i), \quad (8)$$

where

$$\boldsymbol{\mu} = \begin{pmatrix} \beta_0 + \beta_1 \mu_{\xi} + \beta_2 \mu_{\zeta} \\ \mu_{\xi} \\ \mu_{\zeta} \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \beta_1^2 \sigma_{\xi}^2 + \beta_2^2 \sigma_{\zeta}^2 + \tau^2 & \beta_1 \sigma_{\xi}^2 & \beta_2 \sigma_{\zeta}^2 \\ \beta_1 \sigma_{\xi}^2 & \sigma_{\xi}^2 & \sigma_{\xi \zeta} \\ \beta_2 \sigma_{\zeta}^2 & \sigma_{\xi \zeta} & \sigma_{\zeta}^2 \end{pmatrix}.$$

Let $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \mu_{\xi}, \mu_{\zeta}, \tau^2, \sigma_{\xi}^2, \sigma_{\zeta}^2, \sigma_{\xi \zeta})^T$ be the whole parameter vector from model (8). Given the normality assumptions, the likelihood function for $\boldsymbol{\theta}$ has a closed-form expression,

$$L(\boldsymbol{\theta}) \propto \prod_{i=1}^n \{ \det(\boldsymbol{\Sigma} + \Gamma_i) \}^{-\frac{1}{2}} \phi_3 \left[(\boldsymbol{\Sigma} + \Gamma_i)^{-\frac{1}{2}} \left\{ \begin{pmatrix} \hat{\eta}_i \\ \hat{\xi}_i \\ \hat{\zeta}_i \end{pmatrix} - \boldsymbol{\mu} \right\} \right], \quad (9)$$

where ϕ_3 denotes the density function of a three-dimensional standard normal variable. It can be assumed that $s_{\zeta_i}^2$ is known and equal to its estimator $\widehat{var}(\hat{\zeta}_i | \zeta_i)$ provided that the sample size is large enough. However, the within-study variance might be unknown if the information about the additional covariate is given only in terms of the mean values. In this case, it can be replaced by a value that is based on previous studies or expert opinion. As an alternative, consider that subgroup summary information is likely to be available from studies included in the meta-analysis. It can be exploited to derive an approximation of the within-study covariances. The following subsections discuss how to obtain the approximation of the within-study covariances in two cases, namely, when the covariate ζ_i is expressed in terms of log odds and when the covariate ζ_i is the mean value of a characteristic over the studies.

3.2.1 | Approximation of Within-Study Covariances When the Covariate is a Log Odds

Suppose that each study includes information about the number of subjects within each subgroup of events/no events identified by an additional discrete covariate. To fix the ideas, consider a binary indicator Z_i about the number of males/females within the event/no event cases for the treatment group and the control group, as illustrated in Table 1.

The number of events in the treatment group and the control group can be computed as

$$Y_i = Z_{iT1M} + Z_{iT1F}, \quad X_i = Z_{iC1M} + Z_{iC1F}, \quad (10)$$

respectively. Let Z_i denote the number of male patients, and let subgroups be identified by subscripts, so that subscripts M and F indicate female/male, subscripts T and C indicate the treatment group/control group, and subscripts 1 and 0 indicate event/no-event, respectively. For example, Z_{iT1M} denotes the number of treated male patients with the event. Accordingly,

$$Z_i = Z_{iT1M} + Z_{iT0M} + Z_{iC1M} + Z_{iC0M}.$$

Assume that the male indicator at the individual level is distributed as a Bernoulli variable with success probability $\text{expit}(\zeta_i)$,

$$Z_i | \zeta_i \sim \text{Binomial}\{n_i, \text{expit}(\zeta_i)\} \quad (11)$$

where n_i denotes the sample size in study i , $n_i = n_{iT} + n_{iC}$. Accordingly, ζ_i is the unobserved log odds of male, whose observed error-prone version $\hat{\zeta}_i$ and its associated variance $s_{\zeta_i}^2$ are

$$\hat{\zeta}_i = \log \frac{Z_i}{n_i - Z_i}, \quad s_{\zeta_i}^2 = \frac{1}{Z_i} + \frac{1}{n_i - Z_i}, \quad (12)$$

respectively. In this case, it is possible to compute the within-study covariances between the observed risk measures and Z_i . Using the first-order Taylor expansion [20], it can be shown that

$$s_{\eta_i, \zeta_i} \approx \frac{1}{Z_i} \left(\frac{Z_{iT1M}}{Y_i} - \frac{Z_{iT0M}}{n_{iT} - Y_i} \right) - \frac{1}{n_i - Z_i} \left(\frac{Z_{iT1F}}{Y_i} - \frac{Z_{iT0F}}{n_{iT} - Y_i} \right), \quad (13)$$

and

$$s_{\xi_i, \zeta_i} \approx \frac{1}{Z_i} \left(\frac{Z_{iC1M}}{X_i} - \frac{Z_{iC0M}}{n_{iC} - X_i} \right) - \frac{1}{n_i - Z_i} \left(\frac{Z_{iC1F}}{X_i} - \frac{Z_{iC0F}}{n_{iC} - X_i} \right). \quad (14)$$

See the Appendix for details.

TABLE 1 | Subgroup summary information for a binary indicator Z_i .

Treatment group				Control group			
Event		Non-event		Event		Non-event	
Male	Female	Male	Female	Male	Female	Male	Female
Z_{iT1M}	Z_{iT1F}	Z_{iT0M}	Z_{iT0F}	Z_{iC1M}	Z_{iC1F}	Z_{iC0M}	Z_{iC0F}

TABLE 2 | Subgroup summary information when Z_i is the mean value per study.

Treatment group		Control group	
Event	Non-event	Event	Non-event
Z_{iT1}	Z_{iT0}	Z_{iC1}	Z_{iC0}

3.2.2 | An Approximation to Within-Study Covariances When the Covariate is a Mean Value

Suppose that each study includes information about the mean of a covariate Z_i measured on the subjects, subdivided by group (treatment/control) and event (event/no-event). For example, Z_i may represent the mean age of patients in the treatment/control group, subdivided by event/no-event cases. The subgroup summary information can be shown in Table 2, where subscripts T , C , 1, and 0 have the same interpretation as before.

Accordingly,

$$Z_i = \frac{Y_i Z_{iT1} + (n_{iT} - Y_i) Z_{iT0} + X_i Z_{iC1} + (n_{iC} - X_i) Z_{iC0}}{n_i}.$$

Assuming that the unobserved covariate at the individual level follows a Normal distribution with mean ζ_i and variance SD_i^2 , Z_i has a conditional Normal distribution

$$Z_i | \zeta_i \sim N\left(\zeta_i, \frac{SD_i^2}{n_i}\right), \quad (15)$$

so that

$$\hat{\zeta}_i = Z_i, \quad s_{\zeta_i}^2 = \frac{SD_i^2}{n_i}. \quad (16)$$

The covariances between the observed measures of risk and the observed covariate Z_i can be approximated via the first-order Taylor expansion [20], leading to

$$s_{\eta_i, \zeta_i} \approx \frac{1}{n_i}(Z_{iT1} - Z_{iT0}), \quad s_{\xi_i, \zeta_i} \approx \frac{1}{n_i}(Z_{iC1} - Z_{iC0}). \quad (17)$$

See the Appendix for details.

3.2.3 | Pseudo-Likelihood Approach Under Approximate Error Model

The log-likelihood function associated with the model, including the additional covariate information Z_i , whichever the expression of Z_i , can give rise to issues in evaluation and maximization when study-specific information is not available. We mainly refer to the covariance components of the within-study covariance matrix. When available study-specific information is not among those in Tables 1, 2, a different approach is proposed which relies on the concept of pseudo-likelihood [21]. According to this view, we suggest replacing the likelihood function by a pseudo version, developed under the working conditional

independence assumption of $\hat{\eta}_i, \hat{\xi}_i$ and $\hat{\zeta}_i$, which sets to zero the unknown covariances $s_{\eta_i, \zeta_i}, s_{\xi_i, \zeta_i}$. Starting from the approximate measurement error model (7), the pseudo-likelihood function for θ is

$$pL(\theta) \propto \prod_{i=1}^n \left\{ \det(\Sigma + \tilde{\Gamma}_i) \right\}^{-\frac{1}{2}} \phi_3 \left[(\Sigma + \tilde{\Gamma}_i)^{-\frac{1}{2}} \begin{pmatrix} \hat{\eta}_i \\ \hat{\xi}_i \\ \hat{\zeta}_i \end{pmatrix} - \boldsymbol{\mu} \right],$$

where

$$\tilde{\Gamma}_i = \begin{pmatrix} s_{\eta_i}^2 & 0 & 0 \\ 0 & s_{\xi_i}^2 & 0 \\ 0 & 0 & s_{\zeta_i}^2 \end{pmatrix}.$$

In order to account for the risk of misspecification of the likelihood components, the covariance matrix of the estimators can be conveniently computed by using the sandwich formula [22], which is available in the Appendix.

3.2.4 | Pseudo-Likelihood Approach Under Exact Measurement Error Models

As for the classical meta-analysis model with baseline risk information, in the case of study-specific covariates, the exact measurement error model can be specified in place of the approximate error model. A likelihood function can be derived by marginalizing the joint density function of the outcomes, the true measures of risk, and the additional covariate. For example, when the additional covariate is a log odds and subgroup summary information is given in terms of Table 1, the subgroup outcomes can be modeled using multinomial distributions

$$\begin{aligned} & (Z_{iT1M}, Z_{iT0M}, Z_{iT1F}, Z_{iT0F})^\top | (\eta_i, \zeta_i)^\top \\ & \sim \text{Multinomial}_4(p_{iT1M}, p_{iT0M}, p_{iT1F}, p_{iT0F}), \\ & (Z_{iC1M}, Z_{iC0M}, Z_{iC1F}, Z_{iC0F})^\top | (\xi_i, \zeta_i)^\top \\ & \sim \text{Multinomial}_4(p_{iC1M}, p_{iC0M}, p_{iC1F}, p_{iC0F}), \end{aligned} \quad (18)$$

where vectors of probabilities $(p_{iT1M}, p_{iT0M}, p_{iT1F}, p_{iT0F})^\top$ and $(p_{iC1M}, p_{iC0M}, p_{iC1F}, p_{iC0F})^\top$ are functions of $(\eta_i, \zeta_i)^\top$ and $(\xi_i, \zeta_i)^\top$, respectively,

$$\begin{aligned} p_{iT1M} + p_{iT1F} &= \text{expit}(\eta_i) \\ p_{iT1M} + p_{iT0M} &= \text{expit}(\zeta_i) \\ p_{iT1M} + p_{iT0M} + p_{iT1F} + p_{iT0F} &= 1 \\ p_{iC1M} + p_{iC1F} &= \text{expit}(\xi_i) \\ p_{iC1M} + p_{iC0M} &= \text{expit}(\zeta_i) \\ p_{iC1M} + p_{iC0M} + p_{iC1F} + p_{iC0F} &= 1. \end{aligned}$$

In this case, the likelihood function for θ has a more complex expression if compared to the case of the approximate measurement error model, namely,

$$L(\theta) \propto \prod_{i=1}^n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{iT1M}^{Z_{iT1M}} \times p_{iT0M}^{Z_{iT0M}} \times p_{iT1F}^{Z_{iT1F}} \times p_{iT0F}^{Z_{iT0F}} \\ \times p_{iC1M}^{Z_{iC1M}} \times p_{iC0M}^{Z_{iC0M}} \times p_{iC1F}^{Z_{iC1F}} \times p_{iC0F}^{Z_{iC0F}} \times \frac{1}{\sqrt{\tau^2}} \times \frac{1}{\sqrt{\sigma_{\xi}^2}} \times \frac{1}{\sqrt{\sigma_{\zeta}^2}} \\ \times \phi\left(\frac{\eta_i - \beta_0 - \beta_1 \xi_i - \beta_2 \zeta_i}{\tau}\right) \phi\left(\frac{\xi_i - \mu_{\xi}}{\sigma_{\xi}}\right) \phi\left(\frac{\zeta_i - \mu_{\zeta}}{\sigma_{\zeta}}\right) d\eta_i d\xi_i d\zeta_i.$$

Such a likelihood function does not have a closed-form expression and is very expensive to maximize, since every integrand is a product of many functions involving $(\eta_i, \xi_i, \zeta_i)^T$.

When the additional covariate is the mean response of a characteristic and subgroup summary information is given in terms of Table 2, the subgroup outcomes given X_i and Y_i can be modeled using bivariate Normal distributions, that is,

$$\begin{pmatrix} Z_{iT1} \\ Z_{iT0} \end{pmatrix} \Big| \begin{pmatrix} Y_i \\ \zeta_i \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \zeta_i \\ \zeta_i \end{pmatrix}, \begin{pmatrix} \frac{SD_i^2}{Y_i} & \frac{\rho_{iT} SD_i^2}{\sqrt{Y_i(n_{iT}-Y_i)}} \\ \frac{\rho_{iT} SD_i^2}{\sqrt{Y_i(n_{iT}-Y_i)}} & \frac{SD_i^2}{n_{iT}-Y_i} \end{pmatrix} \right), \\ \begin{pmatrix} Z_{iC1} \\ Z_{iC0} \end{pmatrix} \Big| \begin{pmatrix} X_i \\ \zeta_i \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \zeta_i \\ \zeta_i \end{pmatrix}, \begin{pmatrix} \frac{SD_i^2}{X_i} & \frac{\rho_{iC} SD_i^2}{\sqrt{X_i(n_{iC}-X_i)}} \\ \frac{\rho_{iC} SD_i^2}{\sqrt{X_i(n_{iC}-X_i)}} & \frac{SD_i^2}{n_{iC}-X_i} \end{pmatrix} \right) \quad (19)$$

where ρ_{iT} and ρ_{iC} denote within-study correlations which are functions of Y_i and X_i , respectively. Again, the associated likelihood function has a complex form,

$$L(\theta) \propto \prod_{i=1}^n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \phi_2 \\ \left[\begin{pmatrix} \frac{SD_i^2}{Y_i} & \frac{\rho_{iT} SD_i^2}{\sqrt{Y_i(n_{iT}-Y_i)}} \\ \frac{\rho_{iT} SD_i^2}{\sqrt{Y_i(n_{iT}-Y_i)}} & \frac{SD_i^2}{n_{iT}-Y_i} \end{pmatrix}^{-\frac{1}{2}} \left\{ \begin{pmatrix} Z_{iT1} \\ Z_{iT0} \end{pmatrix} - \begin{pmatrix} \zeta_i \\ \zeta_i \end{pmatrix} \right\} \right] \\ \times \phi_2 \left[\begin{pmatrix} \frac{SD_i^2}{X_i} & \frac{\rho_{iC} SD_i^2}{\sqrt{X_i(n_{iC}-X_i)}} \\ \frac{\rho_{iC} SD_i^2}{\sqrt{X_i(n_{iC}-X_i)}} & \frac{SD_i^2}{n_{iC}-X_i} \end{pmatrix}^{-\frac{1}{2}} \left\{ \begin{pmatrix} Z_{iC1} \\ Z_{iC0} \end{pmatrix} - \begin{pmatrix} \zeta_i \\ \zeta_i \end{pmatrix} \right\} \right] \\ \times \{\text{expit}(\eta_i)\}^{Y_i} \{1 - \text{expit}(\eta_i)\}^{n_{iT}-Y_i} \{\text{expit}(\xi_i)\}^{X_i} \\ \{1 - \text{expit}(\xi_i)\}^{n_{iC}-X_i} \\ \times \frac{1}{\sqrt{\tau^2}} \phi\left(\frac{\eta_i - \beta_0 - \beta_1 \xi_i - \beta_2 \zeta_i}{\tau}\right) \frac{1}{\sqrt{\sigma_{\xi}^2}} \phi\left(\frac{\xi_i - \mu_{\xi}}{\sigma_{\xi}}\right) \\ \frac{1}{\sqrt{\sigma_{\zeta}^2}} \phi\left(\frac{\zeta_i - \mu_{\zeta}}{\sigma_{\zeta}}\right) d\eta_i d\xi_i d\zeta_i.$$

When subgroup summary information is unavailable, the construction of the likelihood function can be more cumbersome. While the number of events Y_i and X_i are independent as measured on different subjects, this is not true when Z_i is taken into account, as it typically refers to characteristics of the subjects (Y_i and X_i) included in each study. In addition, a distribution for Z_i given the unobserved ζ_i must be specified as well as a marginal distribution for ζ_i , according to the structural approach

for error correction. Let $f_{Y_i, X_i, Z_i | \eta_i, \xi_i, \zeta_i}$ denote the density function associated with the distribution of $Y_i, X_i, Z_i | \eta_i, \xi_i, \zeta_i$. Let $f_{\eta_i | \xi_i, \zeta_i}$ denote the density function from model (6), and let $f_{\xi_i | \zeta_i}$ and f_{ζ_i} denote the density functions associated with $\xi_i | \zeta_i$ and ζ_i , respectively. An extension of likelihood function (4) to include the uncertainty associated with the additional study-specific covariate Z_i makes the integral three-dimensional, with increasing computational issues,

$$L(\theta) = \prod_{i=1}^n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{Y_i, X_i, Z_i | \eta_i, \xi_i, \zeta_i} f_{\eta_i | \xi_i, \zeta_i} f_{\xi_i | \zeta_i} f_{\zeta_i} d\eta_i d\xi_i d\zeta_i,$$

where the dependence of the functions in the integrand on the parameters of interest is suppressed for convenience of notation.

Deriving the expression of the density functions for Y_i, X_i and Z_i conditional of $(\zeta_i, \xi_i, \zeta_i)^T$ can be difficult. More generally, the evaluation of $L(\theta)$ can be cumbersome given the amount of conditional densities, also in terms of the accuracy measures with the additional covariate component ζ_i . In order to reduce these obstacles, a pseudo-likelihood solution can be adopted, which assumes conditional independence between the observed Y_i and X_i and the covariate Z_i within each study as well as independence between the underlying risk ξ_i and ζ_i . See the Appendix for pseudo-likelihood functions when the covariate is a log odds and when the covariate is a mean value. The evaluation of the integrals requires numerical integration, for example, using Gauss-Hermite quadrature. As for the previous case in Section 3.2.3, the sandwich formula is suggested to compute the covariance matrix of the maximum pseudo-likelihood estimators.

4 | Simulation Study

The likelihood and the pseudo-likelihood approaches have been evaluated in a series of simulation studies covering different scenarios, including additional study-specific covariates. The approaches are compared with the naive analysis, which ignores the presence of measurement errors. Data are generated under the following scenarios:

Scenario 1 Model (5) with an error-free covariate $\zeta_i \sim N(0, 1)$;

Scenario 2 Model (6) with an error-affected covariate, which is a binary indicator of a characteristic as in (18);

Scenario 3 Model (6) with an error-affected covariate, which is the mean response of a characteristic from each study as in (19).

The code for simulation is written in the R programming language [23] and it is available as https://github.com/davidtran2908/control_risk_regression2.

Data are simulated according to the following procedure. First, for a given number n of studies included in the meta-analysis, the number of treated n_{iT} and the number of controls n_{iC} in each study are generated from a uniform distribution $U(15, 200)$. The true risk measures $(\eta_i, \xi_i)^T$ and the study-level information ζ_i

TABLE 3 | Bias, standard error (se), standard deviation (sd) of the maximum likelihood estimators of $(\beta_0, \beta_1, \beta_2, \mu_\xi, \tau^2, \sigma_\xi^2)^\top$ and numbers of convergent solutions over 1000 replicates for the uncorrected approach and the likelihood approach when data follow scenario 1. Underlying risk distributed as a standard Normal.

τ^2	Parameter	Approach	$n = 10$				$n = 20$			
			bias	se	sd	convergence	bias	se	sd	convergence
0.1	β_0	lik	0.007	0.134	0.160	991	-0.006	0.096	0.102	998
		naïve	0.019	0.149	0.176	1000	-0.003	0.104	0.121	1000
	β_1	lik	0.011	0.172	0.214	991	0.012	0.121	0.136	998
		naïve	-0.032	0.164	0.227	1000	-0.061	0.107	0.147	1000
	β_2	lik	-0.014	0.134	0.160	991	-0.010	0.119	0.128	998
		naïve	0.014	0.152	0.180	1000	0.014	0.140	0.157	1000
	μ_ξ	lik	-0.010	0.302	0.314	991	0.016	0.219	0.224	998
	τ^2	lik	-0.056	0.048	0.062	982	-0.036	0.050	0.055	996
		naïve	0.091	0.072	0.125	1000	0.120	0.064	0.090	1000
	σ_ξ^2	lik	-0.088	0.440	0.447	991	-0.072	0.320	0.322	998
0.5	β_0	lik	-0.000	0.232	0.295	998	0.001	0.165	0.182	1000
		naïve	-0.004	0.340	0.405	1000	-0.008	0.179	0.214	1000
	β_1	lik	0.019	0.282	0.363	998	0.012	0.195	0.213	1000
		naïve	-0.027	0.293	0.376	1000	-0.042	0.187	0.218	1000
	β_2	lik	-0.023	0.195	0.247	998	-0.017	0.186	0.204	1000
		naïve	0.014	0.280	0.342	1000	0.023	0.222	0.248	1000
	μ_ξ	lik	-0.002	0.299	0.319	998	-0.006	0.217	0.221	1000
	τ^2	lik	-0.217	0.188	0.233	994	-0.128	0.161	0.171	1000
		naïve	0.114	0.230	0.403	1000	0.094	0.173	0.217	1000
	σ_ξ^2	lik	-0.106	0.434	0.462	998	-0.093	0.315	0.307	1000
1	β_0	lik	0.009	0.296	0.374	999	-0.011	0.251	0.273	1000
		naïve	0.009	0.346	0.432	1000	-0.006	0.258	0.315	1000
	β_1	lik	-0.003	0.381	0.472	999	-0.007	0.254	0.275	1000
		naïve	-0.046	0.408	0.474	1000	-0.058	0.248	0.296	1000
	β_2	lik	-0.012	0.434	0.520	999	-0.009	0.285	0.327	1000
		naïve	0.024	0.455	0.576	1000	0.016	0.311	0.380	1000
	μ_ξ	lik	-0.005	0.293	0.328	999	-0.005	0.218	0.223	1000
	τ^2	lik	-0.365	0.358	0.411	996	-0.251	0.290	0.314	1000
		naïve	-0.035	0.361	0.539	1000	0.066	0.311	0.397	1000
	σ_ξ^2	lik	-0.143	0.417	0.441	999	-0.088	0.317	0.318	1000

are generated from model (5) for scenario 1 and model (6) for scenarios 2 and 3. Values of $(Y_i, X_i)^T$ are then obtained. For scenario 1, these measures are from a binomial distribution (2). For scenario 2, $(Y_i, X_i)^T$ are computed based on formula (10) after the generation of subgroup summary information $(Z_{iT1M}, Z_{iT1F}, Z_{iT0M}, Z_{iT0F}, Z_{iC1M}, Z_{iC1F}, Z_{iC0M}, Z_{iC0F})^\top$, $i = 1, \dots, n$, from model (18), with probabilities p_{iT1M} and p_{iC1M} simulated from uniform distributions $U(0, \min\{\text{expit}(\eta_i), \text{expit}(\zeta_i)\})$ and $U(0, \min\{\text{expit}(\xi_i), \text{expit}(\zeta_i)\})$, respectively. For scenario 3, $(Y_i, X_i)^T$ are derived from a binomial distribution (2) and $(Z_{iT1}, Z_{iT0}, Z_{iC1}, Z_{iC0})^\top$ are generated from model (19), with within-study correlations ρ_{iT} and ρ_{iC} in model (19) generated from a uniform distribution $U(-1, 1)$.

At the final step, the observed error-prone quantities $(\hat{\eta}_i, \hat{\xi}_i)^T$ are obtained using (2) and the associated standard errors are derived using (3). The observed error-prone $\hat{\zeta}_i$ and the associated standard error follow formulas (12) and (16) for scenarios 2 and 3, respectively. Within-study covariances are given as in (13, 14) for scenario 2 and as in (17) for scenario 3.

Values for the underlying risk ξ_i are generated from a standard normal distribution, $\xi_i \sim N(0, 1)$, and from a skew-normal distribution [24], $\xi_i \sim SN(0, 1, -5)$, in order to investigate the performance of the approaches in case of departures from the normality assumptions. Regression parameters are set equal to $(\beta_0, \beta_1, \beta_2)^\top = (0.0, 1.0, 0.8)^\top$. Increasing values of the variance

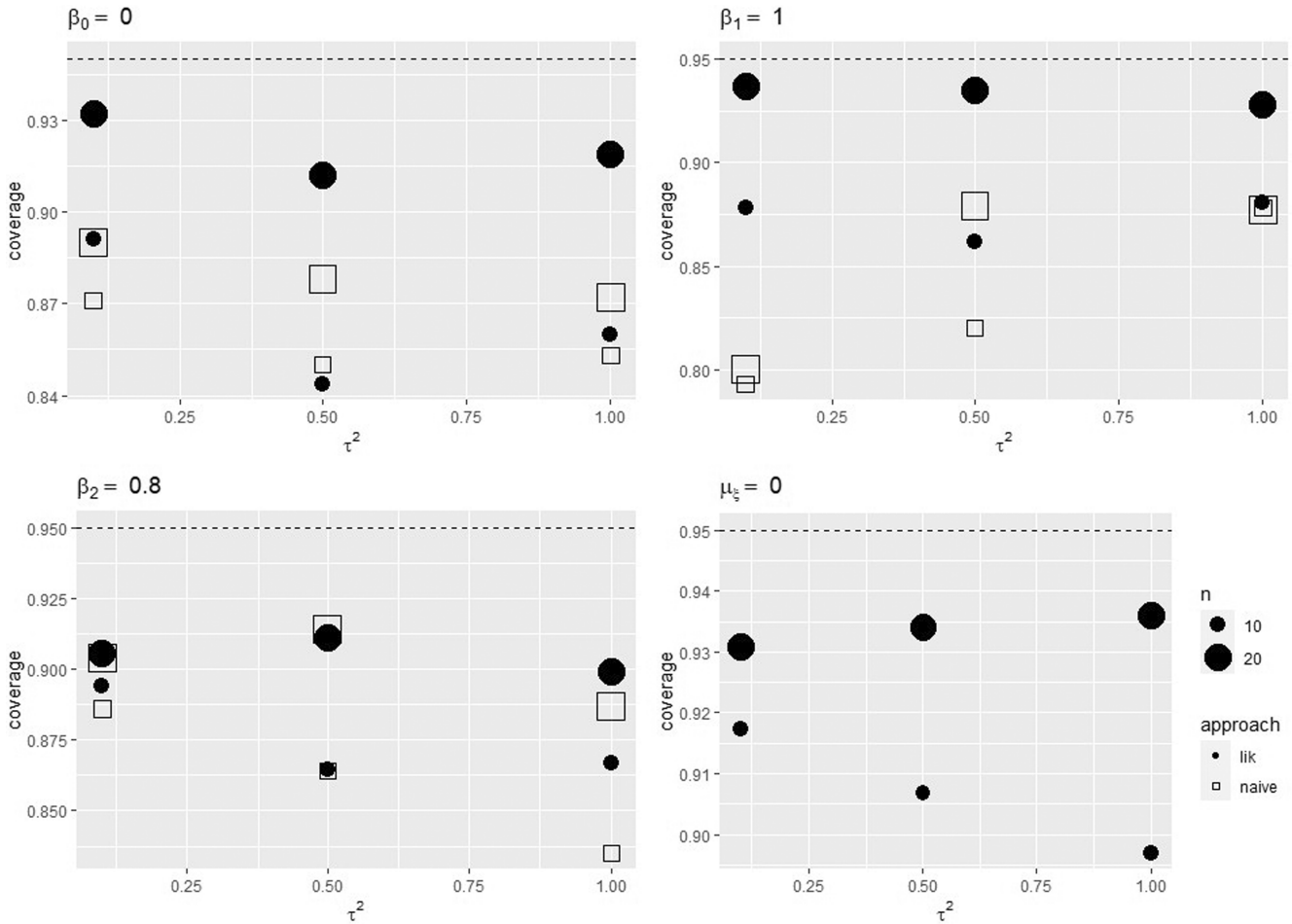


FIGURE 1 | Empirical coverage probabilities of 95% Wald-type confidence intervals for $(\beta_0, \beta_1, \beta_2)^\top$ for the uncorrected approach and the likelihood approach when data follow scenario 1. Underlying risk distributed as a standard Normal.

component is considered, namely, $\tau^2 \in \{0.1, 0.5, 1.0\}$, as well as increasing values of the number of studies included in the meta-analysis, $n \in \{10, 20\}$. One thousand datasets are generated for each scenario and each combination of sample size n , between-study variance τ^2 , and control risk ξ_i distribution. Likelihood maximization, based on the Nelder–Mead algorithm [25], employs the ordinary least squares estimate as a starting value. Numerical integration is performed through Gauss–Hermite quadrature with 10 nodes.

The competing methods are examined in terms of bias, standard error (se), and standard deviation (sd) of the estimator of $(\beta_0, \beta_1, \beta_2, \mu_\xi, \mu_\zeta, \tau^2, \sigma_\xi^2, \sigma_\zeta^2)^\top$, and number of convergent solutions (conv). The empirical coverage probability is computed for 95% confidence intervals of $(\beta_0, \beta_1, \beta_2)^\top$ using the Hessian standard error and the sandwich standard error.

4.1 | Simulation Results

Table 3 reports the bias, the standard error and the standard deviation of the maximum likelihood estimator and the naive estimator of the parameters, when data are generated from model (5) with an additional error-free covariate $\zeta_i \sim N(0, 1)$ and normally distributed control risk. Not surprisingly, results from the

likelihood approach tend to be more satisfactory than the uncorrected solution, with a less biased estimator of the parameter of interest β_1 . The behavior is more evident in the case of increasing values of the between-study variance τ^2 . Figure 1 reports the empirical coverage probabilities of 95% Wald-type confidence intervals for $(\beta_0, \beta_1, \beta_2)^\top$, showing that results from the likelihood approach are closer to the target level than those from the naive analysis, with results improving as the number of studies grows. Similar results are obtained under an error-free additional covariate when moving to a skewed control risk distribution or to model (5) with $\zeta_i \sim \text{Bernoulli}(0.5)$. See Tables S1, S12, and S13 and Figures S1, S4, and S5 in the Supporting Information.

Table 4 and Tables S2 and S3 in Supporting Information report the bias, the standard error, and the standard deviation of the likelihood estimators and the pseudo-likelihood estimators of the parameters in model (6), when data include an error-affected covariate, namely, a binary indicator of a characteristic of the patients, as in model (18). The pseudo-likelihood approach and the likelihood approach based on subgroup summary have similar and satisfactory performance, with almost unbiased estimators. The improvement over the naive analysis in terms of estimate of β_1 becomes relevant when the number of studies increases.

TABLE 4 | Bias, standard error (se), standard deviation (sd) of the maximum likelihood estimators of $(\beta_0, \beta_1, \beta_2, \mu_\xi, \mu_\zeta, \tau^2, \sigma_\xi^2, \sigma_\zeta^2)^\top$ and number of convergent solutions over 1000 replicates for the uncorrected approach, the likelihood approach and the pseudo-likelihood approach when data follow scenario 2 and $\tau^2 = 0.1$. Underlying risk is distributed as a standard Normal.

τ^2	Parameter	Approach	$n = 10$				$n = 20$			
			bias	se	sd	convergence	bias	se	sd	convergence
0.1	β_0	lik	0.000	0.150	0.166	992	0.007	0.102	0.109	997
		pseudo-likelihood	-0.001	0.151	0.167	988	0.007	0.104	0.109	994
		naive	-0.002	0.161	0.184	1000	0.008	0.112	0.125	1000
	β_1	lik	0.004	0.185	0.203	992	0.009	0.126	0.131	997
		pseudo-likelihood	0.006	0.186	0.206	988	0.008	0.126	0.130	994
		naive	-0.036	0.169	0.211	1000	-0.064	0.111	0.139	1000
	β_2	lik	-0.012	0.168	0.197	991	-0.003	0.113	0.124	997
		pseudo-likelihood	-0.012	0.171	0.197	988	-0.001	0.115	0.125	994
		naive	-0.015	0.171	0.211	1000	-0.002	0.116	0.139	1000
	μ_ξ	lik	0.008	0.302	0.309	993	-0.001	0.217	0.225	997
		pseudo-likelihood	0.012	0.303	0.309	988	-0.003	0.219	0.229	994
	μ_ζ	lik	0.002	0.299	0.315	993	0.000	0.215	0.222	997
		pseudo-likelihood	0.004	0.300	0.319	988	0.001	0.216	0.224	994
	τ^2	lik	-0.046	0.060	0.066	979	-0.030	0.054	0.056	987
		pseudo-likelihood	-0.047	0.059	0.066	978	-0.031	0.054	0.058	984
		naive	0.099	0.075	0.124	1000	0.147	0.072	0.107	1000
	σ_ξ^2	lik	-0.090	0.444	0.451	993	-0.089	0.316	0.318	997
		pseudo-likelihood	-0.083	0.446	0.447	988	-0.079	0.320	0.320	994
σ_ζ^2	lik	-0.087	0.428	0.423	993	-0.079	0.305	0.306	997	
	pseudo-likelihood	-0.079	0.431	0.424	988	-0.069	0.308	0.308	994	

Specifically, the estimators of the regression coefficients are nearly unbiased. Although the between-study variance τ^2 is underestimated, the relative bias of its estimator declines in magnitude with the number of studies and the residual variance. The standard errors of the estimators of $(\beta_0, \beta_1, \beta_2, \tau^2)^\top$ reduce when increasing n , as expected, as well as their differences with respect to the associated standard deviations. The empirical coverage probabilities of 95% Wald-type confidence intervals for $(\beta_0, \beta_1, \beta_2)^\top$ from the likelihood-based solutions are comparable (see Figure 2) and notably outperform the ones from the uncorrected approach, whichever the values of the sample size and the between-study variance. For both the likelihood and the pseudo-likelihood approach, the empirical coverage probabilities tend to increase toward the nominal level as n increases and the between-study heterogeneity decreases.

When data include an error-affected covariate defined as a mean value, following model (6) and model (19), the results from the competing approaches are similar to those experienced for data with ζ_i obtained as a binary indicator. See Table 5 for $\tau^2 = 0.1$, Tables S4, S5 in the Supporting Information for $\tau^2 = 0.5$ and $\tau^2 = 1.0$, respectively, and Figure 3.

When control risk follows an asymmetric distribution, $\xi_i \sim SN(0, 1, -5)$, results from the naive analysis are much worse than those under the Normal case, especially when data

are generated from model (6) and model (19). See Tables S6–S8 in the Supporting Material. Improvements given by the likelihood-based solutions are thus more evident, although they also suffer from an increased bias of the estimators of the regression parameters and of the variance components. Bias reduces as the sample size increases. A larger difference between the standard error and standard deviation of the estimators is an expected indication of model misspecification, with evidence for small sample size and large between-study variance, as well as a slightly reduced coverage of confidence intervals; see Figure S2 in the Supporting Information.

The application of the likelihood approach and of the pseudo-likelihood version does not pose computational problems. The failure rate is smaller than 3% in settings characterized by small sample size and small between-study variance, especially when using the likelihood solution in place of the pseudo-likelihood version.

4.2 | A Simulation Study for the Exact Pseudo-Likelihood Approach

The previous simulation study has focused on the approximate measurement error model, mainly for computational convenience. Under an exact measurement error model, in

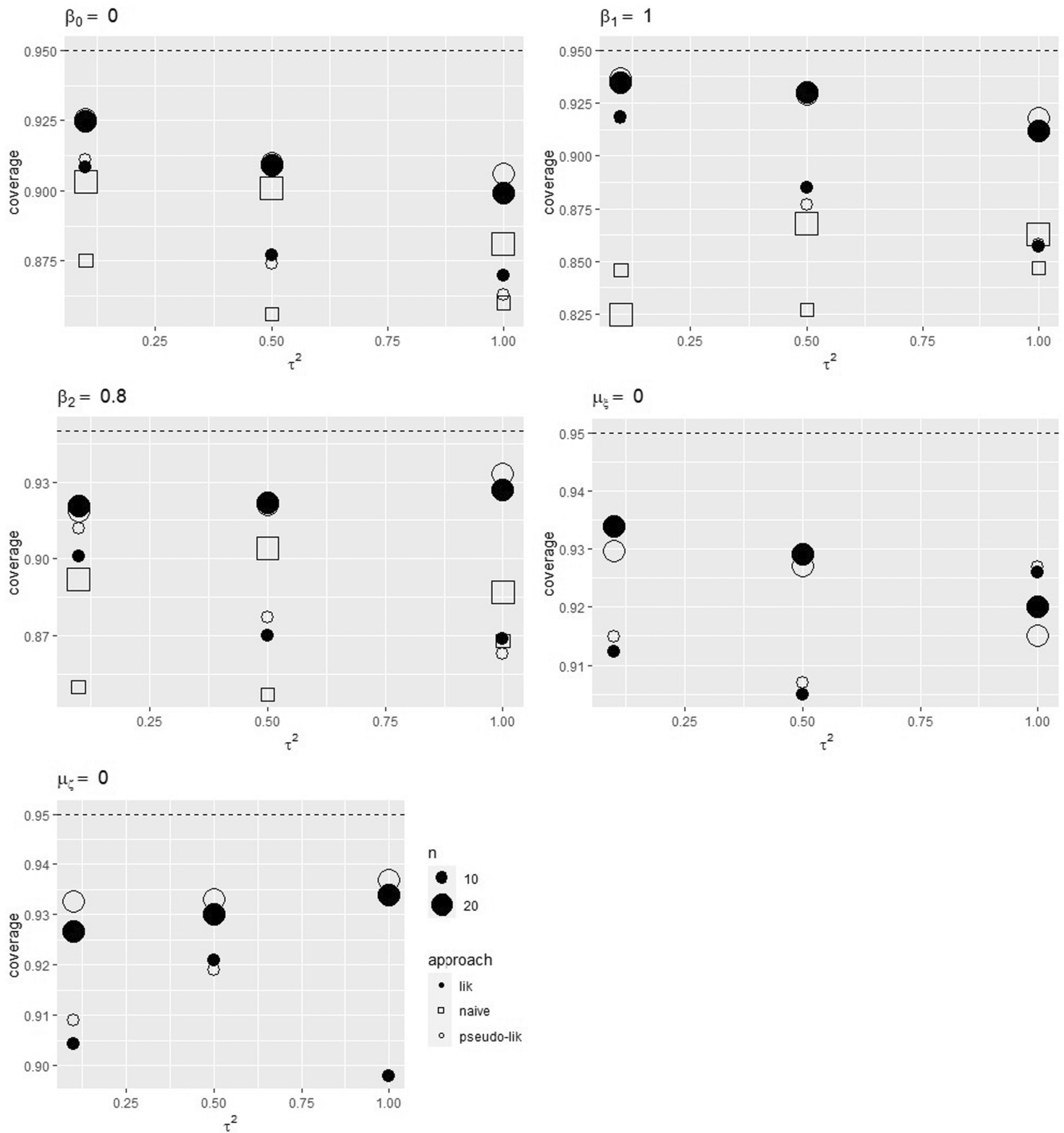


FIGURE 2 | Empirical coverage probabilities of 95% Wald-type confidence intervals for $(\beta_0, \beta_1, \beta_2, \mu_{\zeta}, \mu_{\zeta}^T)^T$ for the uncorrected approach, the likelihood approach, and the pseudo-likelihood approach when data follow scenario 2. Underlying risk distributed as a standard Normal.

fact, the likelihood function is quite complex, as shown in Section 3.2.4. However, in order to investigate the behavior of the pseudo-likelihood approach under an exact measurement error model, a small simulation study has been performed. The simulation study refers to data generated with an additional covariate represented by a binary indicator, as in model (18), with $(n, \tau^2)^T = (10, 0.5)^T$. When evaluating the pseudo-likelihood function using Gaussian–Hermite quadrature, the number of nodes in each dimension is set equal to 10. The simulation is

based on 100 replicates, for computational reasons. Results are reported in Table S14 in the [Supporting Information](#).

These results confirm the previous findings under the approximate measurement error model. The pseudo-likelihood approach provides estimators of the parameters with small bias, especially when focusing on regression coefficients. The slight differences between the standard errors and the associated standard deviations are expected since we are adopting the

TABLE 5 | Bias, standard error (se), standard deviation (sd) of the maximum likelihood estimators of $(\beta_0, \beta_1, \beta_2, \mu_\xi, \mu_\zeta, \tau^2, \sigma_\xi^2, \sigma_\zeta^2)^\top$ and number of convergent solutions over 1000 replicates for the uncorrected approach, the likelihood approach and the pseudo-likelihood approach when data follow scenario 3 and $\tau^2 = 0.1$. Underlying risk distributed as a standard Normal.

τ^2	Parameter	Approach	$n = 10$				$n = 20$			
			bias	se	sd	convergence	bias	se	sd	convergence
0.1	β_0	lik	-0.004	0.148	0.170	995	0.001	0.104	0.118	996
		pseudo-lik	-0.004	0.148	0.169	992	0.001	0.104	0.118	994
		naive	-0.005	0.160	0.185	1000	0.003	0.114	0.134	1000
	β_1	lik	0.008	0.185	0.216	995	0.006	0.129	0.143	996
		pseudo-lik	0.009	0.186	0.216	992	0.007	0.129	0.144	994
		naive	-0.047	0.170	0.225	1000	-0.084	0.111	0.148	1000
	β_2	lik	-0.001	0.165	0.191	995	-0.012	0.111	0.124	996
		pseudo-lik	-0.002	0.165	0.192	992	-0.013	0.111	0.123	994
		naive	0.017	0.176	0.225	1000	0.011	0.120	0.144	1000
μ_ξ	lik	0.005	0.299	0.326	995	-0.001	0.220	0.229	996	
	pseudo-lik	0.005	0.299	0.323	992	-0.001	0.220	0.228	994	
μ_ζ	lik	0.008	0.293	0.322	995	0.003	0.216	0.217	996	
	pseudo-lik	0.006	0.294	0.325	992	0.002	0.216	0.215	994	
τ^2	lik	-0.045	0.059	0.069	985	-0.033	0.054	0.056	984	
	pseudo-lik	-0.046	0.059	0.069	987	-0.034	0.054	0.056	981	
	naive	0.102	0.075	0.135	1000	0.147	0.072	0.100	1000	
σ_ξ^2	lik	-0.109	0.436	0.441	995	-0.080	0.325	0.311	996	
	pseudo-lik	-0.107	0.437	0.450	992	-0.080	0.325	0.316	994	
σ_ζ^2	lik	-0.093	0.411	0.445	995	-0.049	0.304	0.308	996	
	pseudo-lik	-0.092	0.411	0.441	992	-0.051	0.303	0.306	994	

pseudo-likelihood solution. The empirical coverage probabilities associated to the regression coefficients estimators are lower than the nominal level. This gap can be recovered by increasing the number of nodes in the numerical evaluation of integrals involved in the pseudo-likelihood function.

5 | Application: Schizophrenia and COVID-19

Populations with older age, unhealthy lifestyle, physical comorbidities, and psychiatric diseases might be easily affected by COVID-19 [26–28]. The literature indicates that there exist risk factors in schizophrenic patients associated with an increased risk of getting worse impacts of COVID-19 [29], as for example, a dysregulated immune response [30, 31]. Pardamean and colleagues [28] conducted a meta-analysis of ten studies to investigate the relationship between schizophrenia and mortality due to COVID-19. The data, reported in Table 6, include the numbers of deaths (events) and the total number of observed subjects for the group with schizophrenia and the group without schizophrenia. Additional information from each study includes the mean age, the percentage of male patients, and the percentage of diabetic patients. Figure S6 in the Supporting Information reports a forest plot of the data, highlighting a substantial heterogeneity among the studies [34, 39]. Pardamean and colleagues [28] consider a meta-regression model for the risk ratio including several risk factors. Results from the restricted maximum

likelihood approach indicate that schizophrenic patients have a higher risk of mortality due to COVID-19 compared to patients without schizophrenia ($RR = 2.22$; 95% CI: (1.54, 3.20)). Associations between mortality in schizophrenic patients and mean age ($\hat{\beta}_{\text{age}} = -0.033$; 95% CI: (-0.052, -0.015)) and the percentage of smokers ($\hat{\beta}_{\text{smoker}} = 0.027$; 95% CI: (0.008, 0.046)) are statistically significant.

Let the log odds of mortality due to COVID-19 be the risk measure for the group of schizophrenic patients and for the control group. We consider a classical meta-analysis model relating the risk measures, and extend it by considering additional covariates represented by the (standardized) mean age, the log odds of male, and the log odds of diabetes. Other covariates reported in the data are not considered because of the presence of too many missing values. Figure S7 in the Supporting Information reports the scatterplots relating the risk measure for schizophrenic patients to each covariate, suggesting the presence of a linear relationship with the risk in the control group, the mean age, and the log odds of diabetes. The extended version of the basic model with only baseline risk information (1) considers the presence of one additional covariate at a time. Since information about within-study covariances between the additional covariates and the risk measures is unavailable, the pseudo-likelihood approach (6) is adopted. Results are compared with those from the naive analysis.

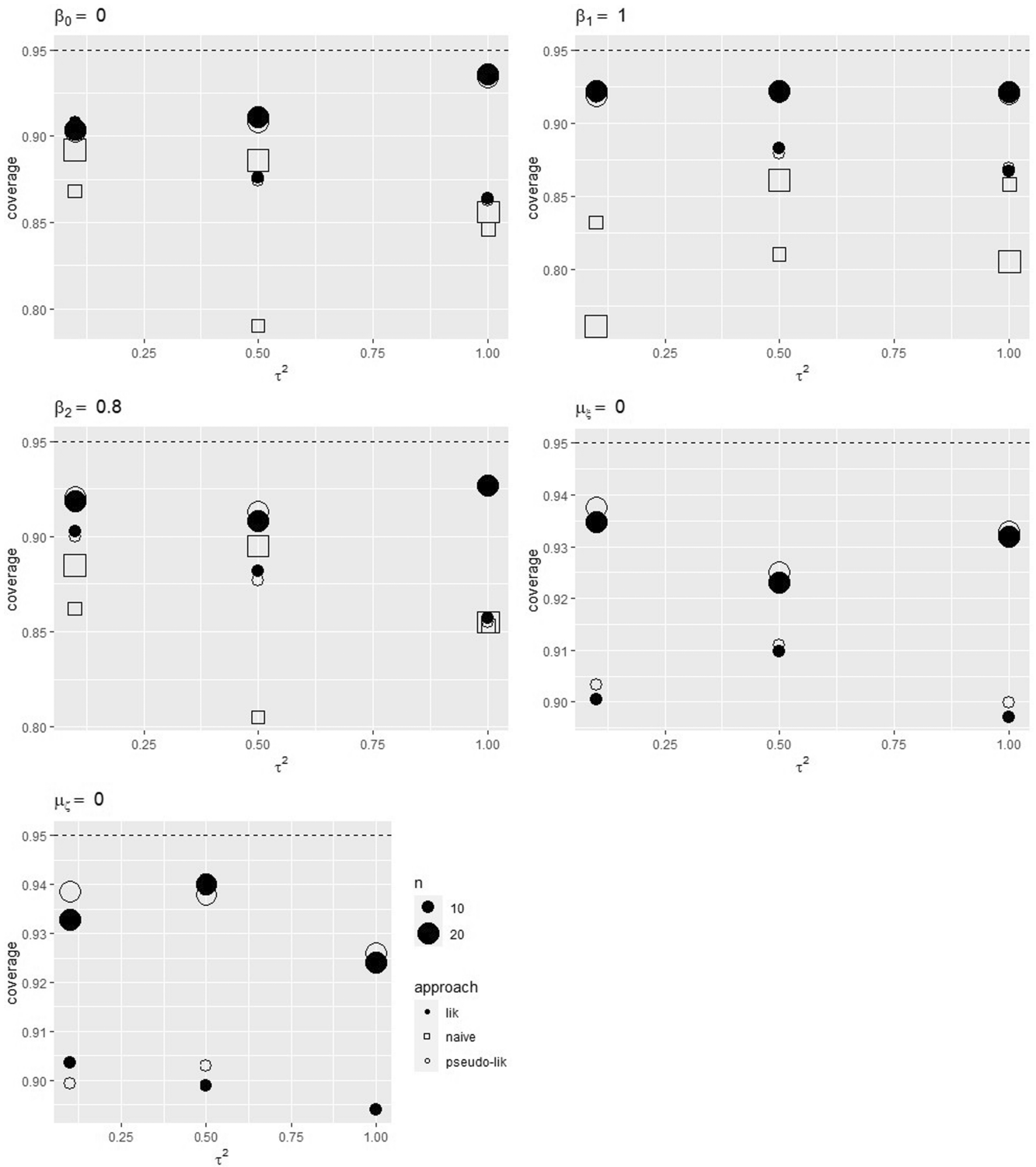


FIGURE 3 | Empirical coverage probabilities of 95% Wald-type confidence intervals for $(\beta_0, \beta_1, \beta_2, \mu_\xi, \mu_\zeta)^\top$ for the uncorrected approach, the likelihood approach, and the pseudo-likelihood approach when data follow scenario 3. Underlying risk distributed as a standard Normal.

Table 7 reports the estimates of the regression coefficients, the residual variance, and the variance for the mismeasured variables obtained from the likelihood approach, the pseudo-likelihood approach, and the naive analysis. According to all the approaches, there is a significant positive association between the measure of risk for the schizophrenic patients and the measure of risk in the control group, either in the presence

or absence of additional covariates. The use of the naive analysis provides smaller values of β_1 than the competing error-corrected approaches, a result which is quite a common effect of the presence of measurement errors in linear regression models. In addition, the naive analysis gives rise to large values of the residual heterogeneity as a consequence of not taking into account the presence of measurement errors. The inclusion of an additional

TABLE 6 | Schizophrenia dataset [28].

Study	With schizophrenia		Without schizophrenia		Characteristics		
	Events	Total	Events	Total	Mean age	% Male	% Diabetes
Barcella et al. [32]	20	984	632	127281	40.3 ± 20.8	48.4	4.6
Tzur Bitan et al. [33]	22	649	7	709	51.5 ± 15.4	60.9	17.1
Fond et al. [34]	4	15	94	1077	63.1 ± 18.5	54.3	23.4
Fond et al. [35]	211	823	10854	49927	70.3 ± 19.2	56.8	27.8
Jeon et al. [36]	6	159	49	2817	55.4 ± 16.2	41.7	15.1
Nemani et al. [37]	20	75	701	6349	54 ± 18.6	47	25.7
Poblador-Plou et al. [38]	11	40	760	4372	67.7 ± 20.7	41.2	11.9
Rivas-Ramirez et al. [39]	2	18	3	69	51.5 ± 14.8	47.1	4.5
Rodriguez-Molinero et al. [40]	2	4	77	414	65.4 ± 16.6	56.9	23.6
Tyson et al. [41]	5	6	70	144	77.6 ± 10.5	50	34

TABLE 7 | Estimates of $(\beta_0, \beta_1, \beta_2, \tau^2)^\top$ and the associated standard errors when using the likelihood approach to fit the classical model with baseline risk information and the pseudo-likelihood approach to fit the model with one additional covariate to schizophrenia dataset [28]. Significant coefficients are highlighted.

Covariate	Approach	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\tau}^2$	$\hat{\sigma}_\xi^2$	$\hat{\sigma}_\zeta^2$
No	likelihood	0.146 (0.329)	0.761 (0.091)	—	0.028 (0.049)	2.459 (1.112)	—
	naive analysis	-0.080 (0.140)	0.709 (0.056)	—	0.527 (0.211)	—	—
Scaled mean age	pseudo-likelihood	0.882 (0.485)	1.071 (0.185)	-0.557 (0.269)	1.6e-05 (5e-04)	2.414 (1.093)	0.922 (0.420)
	naive analysis	0.740 (0.410)	1.004 (0.149)	-0.492 (0.236)	0.466 (0.174)	—	—
Log odds of male	pseudo-likelihood	0.102 (0.186)	0.748 (0.062)	-0.608 (0.417)	0.003 (0.008)	2.514 (1.164)	0.066 (0.031)
	naive analysis	0.034 (0.177)	0.724 (0.057)	-0.422 (0.407)	0.492 (0.184)	—	—
Log odds of diabetes	pseudo-likelihood	0.195 (0.535)	0.743 (0.133)	0.057 (0.274)	0.033 (0.054)	2.578 (1.193)	0.524 (0.255)
	naive analysis	-0.092 (0.195)	0.718 (0.110)	-0.025 (0.251)	0.524 (0.196)	—	—

covariate in the model using a pseudo-likelihood approach gives rise to interesting results. The mean age of the patients is the only resulting significant covariate, with the indication that the risk of mortality decreases with younger subjects. Such a result is in line with the findings in Pardamean and colleagues [28]. When the covariate is significant, its use in the model provides a substantial reduction of the residual variability.

6 | Conclusions

This paper explores the inclusion of additional study-specific covariates other than the control rate in meta-analysis of treatment effectiveness, including baseline risk information, with the aim of better explaining between-study heterogeneity. Inference is performed through a likelihood approach, accounting for the presence of measurement errors associated with the summarized risk measures and covariate information at the study level. When subgroup study-specific summary information associated with additional covariates is available, within-study covariances between the risk measures for cases and controls and the additional covariates are derived using Taylor expansions. When this is not the case, and a fully specified likelihood-based approach cannot be applied, then a pseudo-likelihood solution is exploited. The pseudo-likelihood proposal is defined under a

working conditional independence assumption of the observed quantities, thus avoiding within-study covariances and providing a more robust solution in case of model misspecification with respect to the full likelihood approach. Simulation studies show the applicability of the likelihood and the pseudo-likelihood proposals under a variety of scenarios, including different specifications for the sample size, the between-study heterogeneity, and the underlying risk distribution. Results from both the approaches are satisfactory, although the pseudo-likelihood solution is more appealing in case of deviations from model conditions and it results in less computational effort under an approximate normally distributed measurement error specification.

While the study has been conducted from a inferential frequentist point of view, the use of Bayesian solutions can represent an interesting future extension, especially when the number of additional covariates increases. Specifically, noninformative priors might usefully be assumed for the mean and the variance of the study-specific additional covariates.

Another line of future research includes the application of functional approaches to model the role of baseline risk and additional study-level covariates in meta-analysis of treatment effect, as an alternative to the likelihood solutions. One possibility is represented by SIMEX [12], a simulation extrapolation approach

widely used in the literature to correct for the presence of measurement errors without assumptions on the marginal distribution of the underlying error-prone variables. A first application of SIMEX to model the role of baseline risk is in Guolo [16]. The extension to include additional study-specific covariates, while being of interest for robustness properties, can be challenging as a consequence of computationally intensive simulation procedures.

As suggested by a reviewer, a further extension of the model might consider an interaction between the baseline risk measure and the additional covariate. This extension can be complex, even computationally, due to the interaction that would be created between the measurement errors associated with the variables.

Acknowledgment

Open access publishing facilitated by Università degli Studi di Padova, as part of the Wiley - CRUI-CARE agreement.

Disclosure

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

References

1. M. W. McIntosh, "The Population Risk as an Explanatory Variable in Research Synthesis of Clinical Trials," *Statistics in Medicine* 15, no. 16 (1996): 1713–1728.
2. S. Sharp, S. Thompson, and D. Altman, "The Relation Between Treatment Benefit and Underlying Risk in Meta-Analysis," *BMJ* 313 (1996): 735–738.
3. C. H. Schmid, J. Lau, and M. W. McIntosh, "An Empirical Study of the Effect of the Control Rate as a Predictor of Treatment Efficacy in Meta-Analysis of Clinical Trials," *Statistics in Medicine* 17, no. 17 (1998): 1923–1942.
4. C. H. Schmid, P. C. Stark, J. A. Berlin, P. Landais, and J. Lau, "Meta-Regression Detected Associations Between Heterogeneous Treatment Effects and Study-Level, but Not Patient-Level, Factors," *Journal of Clinical Epidemiology* 57, no. 7 (2004): 683–697.
5. A. Chaimani, "Accounting for Baseline Differences in Meta-Analysis," *Evidence-Based Mental Health* 18 (2015): 23–26.
6. L. R. Arends, A. W. Hoes, J. Lubsen, D. E. Grobbee, and T. Stijnen, "Baseline Risk as Predictor of Treatment Benefit: Three Clinical Meta-Re-Analyses," *Statistics in Medicine* 19, no. 24 (2000): 3497–3518.
7. H. C. van Houwelingen, L. R. Arends, and T. Stijnen, "Advanced Methods in Meta-Analysis: Multivariate Approach and Meta-Regression," *Statistics in Medicine* 21, no. 4 (2002): 589–624.
8. J. P. Higgins and J. A. López-López, "Meta-Regression," in *Handbook of Meta-Analysis*, vol. 7, 1st ed., ed. C. H. Schmid, T. Stijnen, and I. R. White (CRC Press, 2021), 128–149.
9. A. Guolo, "Measurement Error and Misclassification in Meta-Analysis," in *Handbook of Measurement Error Models*, 1st ed., ed. G. Y. Yi, A. Delaigle, and P. Gustafson (Chapman and Hall/CRC, 2021), 481–504.
10. A. Guolo, "Measurement Errors in Control Risk Regression: A Comparison of Correction Techniques," *Statistics in Medicine* 41, no. 1 (2022): 163–179.
11. S. J. Sharp and S. G. Thompson, "Analysing the Relationship Between Treatment Effect and Underlying Risk in Meta-Analysis: Comparison and Development of Approaches," *Statistics in Medicine* 19, no. 23 (2000): 3251–3274.
12. R. J. Carroll, D. Ruppert, L. A. Stefanski, et al., *Measurement Error in Nonlinear Models: A Modern Perspective*, 2nd ed. (Chapman and Hall/CRC, 2006).
13. G. Y. Yi, *Statistical Analysis With Measurement Error or Misclassification: Strategy, Method and Application*, 1st ed. (Springer, 2017).
14. G. Y. Yi, A. Delaigle, and P. Gustafson, *Handbook of Measurement Error Models*, 1st ed. (Chapman and Hall/CRC, 2021).
15. W. Ghidry, T. Stijnen, and H. C. van Houwelingen, "Modelling the Effect of Baseline Risk in Meta-Analysis: A Review From the Perspective of Errors-In-Variables Regression," *Statistical Methods in Medical Research* 22, no. 3 (2013): 307–323.
16. A. Guolo, "The SIMEX Approach to Measurement Error Correction in Meta-Analysis With Baseline Risk as Covariate," *Statistics in Medicine* 33, no. 12 (2014): 2062–2076.
17. Y. Chen, Y. Liu, J. Ning, L. Nie, H. Zhu, and H. Chu, "A Composite Likelihood Method for Bivariate Meta-Analysis in Diagnostic Systematic Reviews," *Statistical Methods in Medical Research* 26 (2017): 914–930.
18. Y. Takwoingi, B. Guo, R. Riley, and J. Deeks, "Performance of Methods for Meta-Analysis of Diagnostic Test Accuracy With Few Studies or Sparse Data," *Statistical Methods in Medical Research* 26 (2017): 1896–1911.
19. A. Guolo, "A Double SIMEX Approach for Bivariate Random-Effects Meta-Analysis of Diagnostic Accuracy Studies," *BMC Medical Research Methodology* 17, no. 1 (2017): 6.
20. P. G. Bagos, "On the Covariance of Two Correlated Log-Odds Ratios," *Statistics in Medicine* 31, no. 14 (2012): 1418–1431.
21. C. Varin, N. Reid, and D. Firth, "Overview of Composite Likelihood Methods," *Statistica Sinica* 21, no. 1 (2011): 5–42.
22. G. Kauermann and R. J. Carroll, "A Note on the Efficiency of Sandwich Covariance Matrix Estimation," *Journal of the American Statistical Association* 96 (2001): 1387–1396.
23. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2024).
24. A. Azzalini, "A Class of Distributions Which Includes the Normal Ones," *Scandinavian Journal of Statistics* 12, no. 2 (1985): 171–178.
25. J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization," *Computer Journal* 7, no. 4 (1965): 308–313.
26. M. S. Gold, D. Sehayek, S. Gabrielli, X. Zhang, C. McCusker, and M. Ben-Shoshan, "COVID-19 and Comorbidities: A Systematic Review and Meta-Analysis," *Postgraduate Medicine* 132, no. 8 (2020): 749–755.
27. T. I. Hariyanto and A. Kurniawan, "Obstructive Sleep Apnea (OSA) and Outcomes From Coronavirus Disease 2019 (COVID-19) Pneumonia: A Systematic Review and Meta-Analysis," *Sleep Medicine* 82 (2021): 47–53.
28. E. Pardamean, W. Roan, K. T. A. Iskandar, R. Prayangga, and T. I. Hariyanto, "Mortality From Coronavirus Disease 2019 (COVID-19) in Patients With Schizophrenia: A Systematic Review, Meta-Analysis and Meta-Regression," *General Hospital Psychiatry* 75 (2022): 61–67.
29. J. Xiong, O. Lipsitz, F. Nasri, et al., "Impact of COVID-19 Pandemic on Mental Health in the General Population: A Systematic Review," *Journal of Affective Disorders* 277 (2020): 55–64.

30. D. Rodrigues-Amorim, T. Rivera-Baltanás, C. Spuch, et al., “Cytokines Dysregulation in Schizophrenia: A Systematic Review of Psychoneuroimmune Relationship,” *Schizophrenia Research* 197 (2018): 19–33.
31. R. A. Kroken, I. E. Sommer, V. M. Steen, I. Dieset, and E. Johnsen, “Constructing the Immune Signature of Schizophrenia for Clinical Use and Research; an Integrative Review Translating Descriptives Into Diagnostics,” *Frontiers in Psychiatry* 9 (2019): 753.
32. C. A. Barcella, C. Polcwiartek, G. H. Mohr, et al., “Severe Mental Illness is Associated With Increased Mortality and Severe Course of COVID-19,” *Acta Psychiatrica Scandinavica* 144, no. 1 (2021): 82–91.
33. D. Tzur Bitan, I. Krieger, K. Kridin, et al., “COVID-19 Prevalence and Mortality Among Schizophrenia Patients: A Large-Scale Retrospective Cohort Study,” *Schizophrenia Bulletin* 47, no. 5 (2021): 1211–1217.
34. G. Fond, V. Pauly, V. Orleans, et al., “Increased In-Hospital Mortality From COVID-19 in Patients With Schizophrenia,” *L'Encéphale* 47, no. 2 (2021): 89–95.
35. G. Fond, V. Pauly, M. Leone, et al., “Disparities in Intensive Care Unit Admission and Mortality Among Patients With Schizophrenia and COVID-19: A National Cohort Study,” *Schizophrenia Bulletin* 47, no. 3 (2021): 624–634.
36. H. L. Jeon, J. S. Kwon, S. H. Park, and J. Y. Shin, “Association of Mental Disorders With SARS-CoV-2 Infection and Severe Health Outcomes: Nationwide Cohort Study,” *British Journal of Psychiatry* 218, no. 6 (2021): 344–351.
37. K. Nemani, C. Li, M. Olfson, et al., “Association of Psychiatric Disorders With Mortality Among Patients With COVID-19,” *JAMA Psychiatry* 78, no. 4 (2021): 380–386.
38. B. Poblador-Plou, J. Carmona-Pírez, and I. Ioakeim-Skoufa, “Baseline Chronic Comorbidity and Mortality in Laboratory-Confirmed COVID-19 Cases: Results From the PRECOVID Study in Spain,” *International Journal of Environmental Research and Public Health* 17, no. 14 (2020): 5171.
39. A. R. Rivas-Ramírez, H. Tendilla-Beltrán, L. E. Gómez-Mendoza, G. Loaiza, and G. Flores, “Patients With Schizophrenia Have Decreased COVID-19 Prevalence Among Hospitalised Patients With Psychiatric and Neurological Diseases: A Retrospective Analysis in Mexican Population,” *International Journal of Clinical Practice* 75, no. 10 (2021): e14528.
40. A. Rodríguez-Molinero, C. Gálvez-Barrón, A. Miñarro, O. Macho, G. F. López, and M. T. Robles, “Association Between COVID-19 Prognosis and Disease Presentation, Comorbidities and Chronic Treatment of Hospitalized Patients,” *PLoS One* 15, no. 10 (2020): e0239571.
41. B. Tyson, L. Erdodi, and A. Shahein, “Predictors of Survival in Older Adults Hospitalized With COVID-19,” *Neurological Sciences* 42 (2021): 3953–3958.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** Supporting Information.

Appendix A

Formula of the Sandwich Covariance Matrix

The sandwich estimate of the covariance matrix of the estimator $\hat{\theta}$ of θ [22] is

$$G(\hat{\theta}) = n^{-1} J^{-1}(\theta) I(\theta) J^{-1}(\theta) \Big|_{\theta=\hat{\theta}},$$

where

$$J(\theta) = n^{-1} E \left\{ \frac{\partial^2 \log pL_i(\theta)}{\partial \theta \partial \theta^T} \right\} \text{ and}$$

$$I(\theta) = n^{-1} E \left[\frac{\partial \log pL_i(\theta)}{\partial \theta} \left\{ \frac{\partial \log pL_i(\theta)}{\partial \theta} \right\}^T \right],$$

$pL_i(\theta)$ is the contribution of study i to the pseudo-likelihood for θ , $\partial f(x)/\partial x$ is the derivative of $f(x)$ with respect to x and $E(X)$ is the expectation of X . Matrices $J(\theta)$ and $I(\theta)$ can be consistently estimated by their empirical counterpart.

Appendix B

Pseudo-Likelihood Approach Under Exact Measurement Error Models

If the covariate is a log odds and under normality assumptions for η_i, ξ_i, ζ_i , a pseudo-likelihood function can be derived using model (11), as

$$pL(\theta) \propto \prod_{i=1}^n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \{\text{expit}(\eta_i)\}^{Y_i} \{1 - \text{expit}(\eta_i)\}^{n_i - Y_i} \\ \times \{\text{expit}(\xi_i)\}^{X_i} \{1 - \text{expit}(\xi_i)\}^{n_i - X_i} \{\text{expit}(\zeta_i)\}^{Z_i} \\ \{1 - \text{expit}(\zeta_i)\}^{n_i - Z_i} \\ \times \frac{1}{\sqrt{\tau^2}} \phi\left(\frac{\eta_i - \beta_0 - \beta_1 \xi_i - \beta_2 \zeta_i}{\tau}\right) \frac{1}{\sqrt{\sigma_\xi^2}} \phi\left(\frac{\xi_i - \mu_\xi}{\sigma_\xi}\right) \\ \frac{1}{\sqrt{\sigma_\zeta^2}} \phi\left(\frac{\zeta_i - \mu_\zeta}{\sigma_\zeta}\right) d\eta_i d\xi_i d\zeta_i.$$

When the covariate is a mean value, a pseudo-likelihood function can be derived using model (15), as

$$pL(\theta) \propto \prod_{i=1}^n \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \{\text{expit}(\eta_i)\}^{Y_i} \{1 - \text{expit}(\eta_i)\}^{n_i - Y_i} \\ \times \{\text{expit}(\xi_i)\}^{X_i} \{1 - \text{expit}(\xi_i)\}^{n_i - X_i} \\ \times \phi\left(\frac{Z_i - \zeta_i}{SD_i/\sqrt{n_i}}\right) \times \frac{1}{\sqrt{\tau^2}} \phi\left(\frac{\eta_i - \beta_0 - \beta_1 \xi_i - \beta_2 \zeta_i}{\tau}\right) \\ \times \frac{1}{\sqrt{\sigma_\xi^2}} \phi\left(\frac{\xi_i - \mu_\xi}{\sigma_\xi}\right) \frac{1}{\sqrt{\sigma_\zeta^2}} \phi\left(\frac{\zeta_i - \mu_\zeta}{\sigma_\zeta}\right) d\eta_i d\xi_i d\zeta_i.$$

Appendix C

Proof of Formula of s_{η_i, ζ_i}

Case With ζ_i as a Log Odds

The within-study covariance between treatment risk and an additional covariate express a log odds can be estimated using the first-order Taylor's expansion around the means [20]:

$$s_{\eta_i, \zeta_i} = \text{cov}(\hat{\eta}_i, \hat{\zeta}_i) \\ = \text{cov}\left(\log \frac{Y_i}{n_{iT} - Y_i}, \log \frac{Z_i}{n_i - Z_i}\right) \\ = \text{cov}(\log Y_i, \log Z_i) - \text{cov}\{\log(n_{iT} - Y_i), \log Z_i\} \\ - \text{cov}\{\log Y_i, \log(n_i - Z_i)\} + \text{cov}\{\log(n_{iT} - Y_i), \\ \log(n_i - Z_i)\} \\ \approx \frac{1}{E(Y_i)E(Z_i)} \text{cov}(Y_i, Z_i) \\ - \frac{1}{E(n_{iT} - Y_i)E(Z_i)} \text{cov}(n_{iT} - Y_i, Z_i) \\ - \frac{1}{E(Y_i)E(n_i - Z_i)} \text{cov}(Y_i, n_i - Z_i) \\ + \frac{1}{E(n_{iT} - Y_i)E(n_i - Z_i)} \text{cov}(n_{iT} - Y_i, n_i - Z_i),$$

where cov denotes “within-study covariance” or “covariance” given $(\eta_i, \xi_i, \zeta_i)^\top$. In the above expansion, the expected values of Y_i and Z_i can be estimated with their observed values. Let U_{ik} and V_{ik} , $k = 1, \dots, n$, denote the indicator of the event and the indicator of the male of subject k in the treatment group of study i , respectively. The within-study covariance between Y_i and Z_i can be estimated as

$$\begin{aligned} cov(Y_i, Z_i) &= cov(Y_i, Z_{iT1M} + Z_{iT0M} + Z_{iC1M} + Z_{iC0M}) \\ &= cov(Y_i, Z_{iT1M} + Z_{iT0M}) \\ &= \sum_{j=1}^{n_{iT}} \sum_{k=1}^{n_{iT}} cov(U_{ij}, V_{ik}) \\ &= \sum_{k=1}^{n_{iT}} cov(U_{ik}, V_{ik}) \\ &= n_{iT} cov(U_{i1}, V_{i1}) \\ &= n_{iT} \{E(U_{i1}V_{i1}) - E(U_{i1})E(V_{i1})\} \\ &= n_{iT} \{Pr(\text{a treated patient is male and with event}) \\ &\quad - Pr(\text{a treated patient is with event}) \\ &\quad \quad Pr(\text{a treated patient is male})\} \\ &\approx n_{iT} \left\{ \frac{Z_{iT1M}}{n_{iT}} - \frac{Y_i}{n_{iT}} \times \frac{(Z_{iT1M} + Z_{iT0M})}{n_{iT}} \right\} \\ &= Z_{iT1M} - \frac{Y_i(Z_{iT1M} + Z_{iT0M})}{n_{iT}}. \end{aligned}$$

Other within-study covariances in the expansion of s_{η_i, ζ_i} can be estimated similarly,

$$\begin{aligned} cov(n_{iT} - Y_i, Z_i) &\approx Z_{iT0M} - \frac{(n_{iT} - Y_i)(Z_{iT1M} + Z_{iT0M})}{n_{iT}}, \\ cov(Y_i, n_i - Z_i) &\approx Z_{iT1F} - \frac{Y_i(Z_{iT1F} + Z_{iT0F})}{n_{iT}}, \\ cov(n_{iT} - Y_i, n_i - Z_i) &\approx Z_{iT0F} - \frac{(n_{iT} - Y_i)(Z_{iT1F} + Z_{iT0F})}{n_{iT}}. \end{aligned}$$

Hence, the within-study covariance between $\hat{\eta}_i$ and $\hat{\zeta}_i$ is

$$s_{\eta_i, \zeta_i} \approx \frac{1}{Z_i} \left(\frac{Z_{iT1M}}{Y_i} - \frac{Z_{iT0M}}{n_{iT} - Y_i} \right) - \frac{1}{n_i - Z_i} \left(\frac{Z_{iT1F}}{Y_i} - \frac{Z_{iT0F}}{n_{iT} - Y_i} \right).$$

Case With ζ_i as a Mean Value

The within-study covariance between the treatment risk and an additional covariate expressed as a mean value can be approximated using the first-order Taylor’s expansion around the means [20]. Let Z_{iT} and Z_{iC} denote mean values in the treatment group and in the control group, respectively. Thus,

$$\begin{aligned} s_{\eta_i, \zeta_i} &= cov(\hat{\eta}_i, \hat{\zeta}_i) \\ &= cov\left(\log \frac{Y_i}{n_{iT} - Y_i}, Z_i\right) \\ &= cov\left(\log \frac{Y_i}{n_{iT} - Y_i}, \frac{n_{iT}Z_{iT} + n_{iC}Z_{iC}}{n_i}\right) \\ &= cov\left\{\log Y_i - \log(n_{iT} - Y_i), \frac{n_{iT}}{n_i} Z_{iT}\right\} \\ &= \frac{n_{iT}}{n_i} cov(\log Y_i, Z_{iT}) - \frac{n_{iT}}{n_i} cov\{\log(n_{iT} - Y_i), Z_{iT}\} \\ &\approx \frac{n_{iT}}{n_i} \frac{1}{E(Y_i)} cov(Y_i, Z_{iT}) \\ &\quad - \frac{n_{iT}}{n_i} \frac{1}{E(n_{iT} - Y_i)} cov(n_{iT} - Y_i, Z_{iT}). \end{aligned}$$

The expected value of Y_i can be estimated with its observed value. Let U_{ik} denote the indicator of the event and V_{ik} denote the age of subject k in the treatment group of study i . The covariance between Y_i and Z_{iT} can be estimated as

$$\begin{aligned} cov(Y_i, Z_{iT}) &= \frac{1}{n_{iT}} cov\left(\sum_{j=1}^{n_{iT}} U_{ij}, \sum_{k=1}^{n_{iT}} V_{ik}\right) \\ &= \frac{1}{n_{iT}} \sum_{j=1}^{n_{iT}} \sum_{k=1}^{n_{iT}} cov(U_{ij}, V_{ik}) \\ &= \frac{1}{n_{iT}} \sum_{k=1}^{n_{iT}} cov(U_{ik}, V_{ik}) \\ &= \frac{1}{n_{iT}} n_{iT} cov(U_{i1}, V_{i1}) \\ &= E(U_{i1}V_{i1}) - E(U_{i1})E(V_{i1}). \end{aligned}$$

Since $U_{i1}V_{i1}$ is equal to V_{i1} if treated patient 1 is with event and zero otherwise, its expected value can be estimated by $Y_i Z_{iT1} / n_{iT}$, so that

$$cov(Y_i, Z_{iT}) \approx \frac{Y_i}{n_{iT}} (Z_{iT1} - Z_{iT}).$$

The within-study covariance between $n_{iT} - Y_i$ and Z_{iT} can be estimated similarly,

$$cov(n_{iT} - Y_i, Z_{iT}) \approx \frac{n_{iT} - Y_i}{n_{iT}} (Z_{iT0} - Z_{iT}).$$

Therefore, the within-study covariance between $\hat{\eta}_i$ and $\hat{\zeta}_i$ is

$$s_{\eta_i, \zeta_i} \approx \frac{1}{n_i} (Z_{iT1} - Z_{iT0}).$$