Published in final edited form as:

Stat Med. 2013 February 28; 32(5): 772–786. doi:10.1002/sim.5552.

Design considerations for case series models with exposure onset measurement error

Sandra M. Mohammed^a, Lorien S. Dalrymple^b, Damla entürk^c, and Danh V. Nguyen^{a,*†}

^aDivision of Biostatistics, Department of Public Health Sciences, University of California, Davis, California 95616, U.S.A

^bDivision of Nephrology, Department of Medicine, University of California, Sacramento, California 95691, U.S.A

^cDepartment of Biostatistics, University of California, Los Angeles, California 90095, U.S.A

Summary

The case series model allows for estimation of the relative incidence of events, such as cardiovascular events, within a pre-specified time window after an exposure, such as an infection. The method requires only cases (individuals with events) and controls for all fixed/time-invariant confounders. The measurement error case series model extends the original case series model to handle imperfect data, where the timing of an infection (exposure) is not known precisely. In this work, we propose a method for power/sample size determination for the measurement error case series model. Extensive simulation studies are used to assess the accuracy of the proposed sample size formulas. We also examine the magnitude of the relative loss of power due to exposure onset measurement error, compared to the ideal situation where the time of exposure is measured precisely. To facilitate the design of case series studies, we provide publicly available web-based tools for determining power/sample size for both the measurement error case series model as well as the standard case series model.

Keywords

case series models; exposure timing measurement error; longitudinal observational database; non-homogeneous Poisson process; sample size

1 Introduction

The case series (CS) model (also called self-controlled case series) was originally developed by Farrington [1] to assess the relationship between adverse events following transient exposures, such as vaccination over time [2]. More precisely, it is a method for estimating the relative incidence of events in a pre-specified time period of interest after an exposure. Since its introduction, the CS method has been applied in a variety of epidemiological/ biomedical studies, including assessment of the association between the use of prescription

*Correspondence to: Danh V. Nguyen, Division of Biostatistics, Department of Public Health Sciences, University of California, Davis, CA 95616.

†ucdnguyen@ucdavis.edu

medications and the risk of motor vehicle crashes [3], and the risk of myocardial infarction and stroke after acute infection in the general population [4]. More recently, Dalrymple *et al.* [5] examined the risk of cardiovascular events after infection-related hospitalizations in the older U.S. dialysis population using the United States Renal Data System (USRDS). For patients on dialysis, infection and cardiovascular disease are the leading cause of hospitalization and death [6].

The CS model is derived from a non-homogeneous Poisson cohort model conditioned on the number of events per individual to be at least one (and an individual's exposure history), and requires only cases, i.e., individuals with one or more events. An individual's observation period is essentially divided into one or more risk periods, specified *a priori*, and control (non-risk) time periods. For example, in the aforementioned USRDS application, it is of interest to estimate the relative incidence of cardiovascular events, such as myocardial infarction, during the 30-day risk period following an infection in patients on dialysis. As detailed in the excellent expository papers by Whitaker *et al.* [7, 8], there are several notable advantages of the CS model. First, it provides straightforward estimation of, and valid inference about, the incidence of events in the risk periods relative to the control period based on cases only. Secondly, it controls for all fixed or time-invariant confounders, such as co-existing illness or other covariates not easily measured in epidemiological or observational studies. For example, in the dialysis cohort from the USRDS, dialysis patients who do and do not acquire infections likely differ in important ways not easily measured or even understood. Third, the model may incorporate age-variation (or time-variation) in the baseline incidence rates. We note that the exposure history for each individual is assumed to be precisely known, i.e., the *dates/times* when exposures occurred are assumed to be measured without error.

Recently, Mohammed *et al.* [9] proposed the measurement error case series (MECS) models, which extend the case series models to account for error in the exact date or time of exposure onset. We refer to this as exposure onset measurement error (i.e., the error in the timing of exposure), which is distinct from measurement error in the covariates (which has been extensively studied; e.g., see [10]). This proposal was motivated by the investigation of an infection-cardiovascular risk association in patients on dialysis using data from the USRDS, where the exact date of infection (exposure) onset cannot be ascertained (e.g., based on hospital claims data). Under this limitation, the discharge date was used as a surrogate marker for the time of infection as it reasonably assures that the infection had occurred by this date. Thus, a positive additive exposure onset measurement error model was proposed: $w = v + u$, where w is the observed exposure onset time (infection-related discharge time), v is the true (unobserved) exposure onset time and u is a positive error in the timing of the infection with mean $\mu_u = E(u)$. For example, assuming that infection is equally likely during the hospitalization stay, Mohammed *et al.* [9] used the length of hospitalization to obtain the estimate $\hat{\mu}_u = 5.5$ days. Thus, on average, infection occurs about 6 days prior to infection-related hospitalization discharge. We note here that the relative incidence estimates will be biased generally when ignoring measurement error. Background and necessary aspects of the MECS model are reviewed in Section 2.

We note that the exposure onset measurement error, as illustrated by the example of assessing infection and cardiovascular events using the USRDS longitudinal database, cannot be avoided at the planning stage. This is relevant to other applications of the case series method to longitudinal observational databases (e.g., hospital claims or administrative databases) such as adverse events due to medications or other types of hospitalizations. Although the error cannot be avoided by design, one can plan for the additional sample size needed to detect a hypothesized effect size.

There are several specific aims of this paper. The first is to derive sample size formulas for the MECS model (with and without age effects). By determining the true target of the naive (biased) estimates, we describe a simple way to utilize an existing sample size formula for the case series model proposed by Musonda *et al.* [11] to determine sample size for the corresponding MECS model. The second aim is to examine the relative loss of power to detect the true relative incidence of events attributed to exposure onset measurement error under the the MECS model relative to the CS model (without exposure onset measurement error). For this we focus on the naive hypothesis testing of no effect (i.e., relative incidence equals one) based on the naive estimate that ignores exposure onset measurement error. This approach is similar to classical measurement error in the covariates (e.g., see [10]) since testing based on biased-corrected estimators, which generally have substantially higher variance relative to naive estimators that ignore measurement error, results in a dramatic loss of power. Understanding the loss of power is informative to researchers considering case series studies with additive exposure onset measurement error as described above. We illustrate the (asymptotic) validity of the naive test, i.e., its Type I error rate approaches the nominal test level. Finally, we provide a publicly available web-based tool for determining sample size (or power) in planning case series studies, both with and without exposure onset measurement error.

The paper is organized as follows. In Section 2, we provide the needed background/preliminary materials, including the CS and MECS models and the existing sample size calculation method for the CS model under the assumption of precisely measured exposure times. The proposed method to determine sample size for study planning and to study the relative loss of power due to exposure onset measurement error is described in Section 3. The simulation studies for MECS models with and without age effects are presented in Section 4. Here we illustrate the magnitude of the loss of power for testing the null hypothesis of no effect as the average amount of measurement error, μ_u , increases. Assessment of the accuracy of the sample size formulas for the MECS model using simulations is also summarized in Section 4. Section 5 illustrates the proposed sample size formula using infection-related hospitalizations and cardiovascular events in the dialysis population, as well as a publicly available web-based application for sample size planning. We conclude with a brief discussion in Section 6.

2 Models and preliminaries

2.1 The CS and MECS models

In this section we provide the necessary background on the case series model [1] and the measurement error case series model, i.e., the case series model in the presence of exposure

onset measurement/timing error [9]. The case series model compares the incidence of events within a risk period of interest relative to the incidence in the baseline period, within each individual. Given the exposure history over the observation period for individual i , the number of events in each age-risk interval, denoted n_{ijk} , is assumed to follow a non-homogeneous Poisson process with rate $\lambda_{ijk} = \exp(\phi_i + \lambda_j + \beta k)$, i.e., $n_{ijk} \sim \text{Poisson}(e_{ijk}\lambda_{ijk})$, where e_{ijk} is the length of time in the j^{th} age group and k^{th} risk period for individual i . Here the parameters ϕ_i , δ_j and β are, respectively, the individual-specific, j^{th} age group (relative to age group $j = 1$) and risk group (relative to baseline period $k = 0$) effects, with $\delta_1 = 0$. The main parameter of interest is β , the log relative incidence of events in the exposure risk period.

Farrington [1] showed that when conditioned on $n_{i..} = \sum_{jk} n_{ijk} - 1$, where $n_{i..}$ is the total number of events for individual i , the kernel of the case series (conditional) likelihood is product multinomial. More specifically, the contribution to the likelihood from subject i is given as $L_i(\delta_1, \dots, \delta_J, \beta) = \prod_{j,k} \pi_{ijk}^{n_{ijk}}$, with probabilities

$$\pi_{ijk} = \frac{e_{ijk}\lambda_{ijk}}{\sum_{s=1}^J \sum_{t=0}^1 e_{ist}\lambda_{ist}} = \frac{e_{ijk} \exp(\delta_j + \beta k)}{\sum_{s=1}^J \sum_{t=0}^1 e_{ist} \exp(\delta_s + \beta t)} \quad (1)$$

The term “self-controlled” refers to the fact that the individual effects ϕ_i cancel out, thus, self-controlling for all fixed covariates.

When the precise time of exposure (e.g., infection) is unknown, Mohammed *et al.* [9] proposed the measurement error case series model to account for this uncertainty. The MECS model was motivated by the use of the hospitalization data from the USRDS to determine whether there is an increase in the relative incidence of cardiovascular events (e.g., myocardial infarction, unstable angina, stroke or transient ischemic attack) within a specified window of time after an infection (e.g., 30 days after an infection). Since patient hospitalization records rely on discharge diagnoses, the available time of an infection-related hospitalization discharge is used in place of the true (unknown) time of infection, which occurs prior to the discharge time. Therefore, a positive additive exposure onset measurement/timing error model,

$$w_i = v_i + u_i, \quad (2)$$

was proposed, where w_i is the observed exposure onset time (infection-related discharge time), v_i is the true (unobserved) exposure onset time and u_i is a positive measurement error ($u_i > 0$) with mean $\mu_u = E(u_i) > 0$. The CS model combined with the exposure timing error model (2) is called the MECS model. Note that the amount of measurement error in the exposure time, u_i , cannot be unrestricted and a practical and necessary assumption is that u_i is less than the length of the risk period of interest. For a concrete example, consider the relative incidence of events associated with the 30-day risk period after an infection of interest. Then the uncertainty in the time when the infection actually occurred should not exceed 30 days; otherwise, one could not estimate the relative incidence in the 30-day risk period after an infection because $u_i > 30$ amounts to not having *any* reliable data for

estimation. As indicated by Mohammed *et al.* [9], this practical assumption on the magnitude of the measurement error essentially ensures that there must be some amount of reliable data for estimation.

As we will consider the naive test of the hypothesis of no effect for sample size determination in Section 3 below, we denote the naive (conditional) MLE of β (ignoring exposure onset measurement error) by $\hat{\beta}^*$, which is obtained from solving the set of likelihood equations

$$\begin{aligned} N^{-1} \sum_{i=1}^N \sum_{j=1}^J \left(\tilde{n}_{ij1} - n_{i..} \hat{\pi}_{ij1}^* \right) &= 0, \\ N^{-1} \sum_{i=1}^N \sum_{k=0}^1 \left(\tilde{n}_{ijk} - n_{i..} \hat{\pi}_{ijk}^* \right) &= 0, \quad j=2, \dots, J, \end{aligned}$$

where $\hat{\pi}_{ijk}^* = e_{ijk} \exp(\hat{\delta}_j^* + \hat{\beta}^* k) \sum_{s=1}^J \sum_{t=0}^1 \exp(\hat{\delta}_s^* + \hat{\beta}^* t)$, $n_{i..}$ is the total number of events for subject i , and $\tilde{n}_{ijk}$ is the number of events in age group j and risk group k based on the observed exposure times $\{w_i\}$ and interval lengths $\{e_{ijk}\}$.

Following the work of Musonda *et al.* [11] on sample size determination for the case series model, for the purpose of study planning, we assume that individuals exposed have a risk period of length e_1 (e.g., 30 days after an infection) and that each individual's follow-up observation period is of the same duration. In the simpler case of no age effects, denote the length of observation period to be $e_0 + e_1$, where e_0 is the length of the baseline/control period and also let $r = e_1/(e_1 + e_0)$ be the proportion of the risk period to the total observation period. In the study design phase, e_0 and e_1 are to be specified by the investigator. The proportion of individuals in the population exposed is denoted by p .

Musonda *et al.* [11] found that age effects can have a large effect on study power. Thus, to model age effects, the total observation period is partitioned into J age groups of length

$e_j = \sum_{k=0}^1 e_{ijk}$, $j=1, \dots, J$. (Note that it is assumed that the lengths of the age groups are the same across individuals.) As in the case of no age groups, the length of the risk period is e_1 . More precisely, $e_{ij1} = e_1$ if the exposure occurs in age group j and it is zero in the remaining age groups ($l \neq j$; $l = 1, \dots, J$). Also, let p_j be the proportion of individuals in the

population exposed in the j^{th} age group, $j = 1 \dots, J$, where $p_0 = 1 - \sum_{j=1}^J p_j$ is the proportion of individuals unexposed. Further assume that the risk period is fully contained within an age group (i.e., $e_1 < e_j$), which simplifies the calculations considerably; see Musonda *et al.* [11]. In the case of no age effect, subscript j can be dropped.

2.2 Existing sample size method when exposure is measured without error

Sample size calculation formulas for the CS models, both with and without age effects as described above, have been developed by Musonda *et al.* [11]. The sample size needed to achieve a desired power for testing the null hypothesis that the relative incidence of interest $\rho \equiv \exp(\beta)$ equals one, i.e., $H_0: \beta = 0$ versus the alternative $H_1: \beta \neq 0$, at a fixed significance

level can be based on the signed root likelihood ratio test, which has an asymptotic normal distribution.

First, for the CS model without age effects, let n be the total number of events across all individuals, let n_1 be the total number of events in exposed individuals only and let x be the total number of events occurring in the risk period. The test statistic [11] is

$$T = \text{sgn}(\hat{\beta}) \sqrt{2} \left[x\hat{\beta} - n_1 \log(\hat{\rho} + 1 - r) \right]^{1/2},$$

where $\hat{\beta}$ denotes the maximum likelihood estimate, $\hat{\rho} = \exp(\hat{\beta})$ and $r = e_1/(e_1 + e_0)$, as defined earlier in Section 2.1. Under H_0 , $T \approx N(0, 1)$ and under H_1 , $T \approx N(\text{sgn}(\beta) \sqrt{nA}, B)$ with

$$A = \frac{2p\tilde{\rho}}{1 + pr(\rho - 1)} \left[\frac{pr\beta}{\tilde{\rho}} - \log(\tilde{\rho}) \right] \quad \text{and} \quad B = \frac{\beta^2}{A} \frac{p}{1 + pr(\rho - 1)} \frac{pr(1 - r)}{\tilde{\rho}}, \quad (4)$$

where $\tilde{\rho} = pr + 1 - r$. Thus, the sample size (i.e., the number of events), needed to achieve 100 γ percent power at level of significance α is given by

$$n = \left(z_{1-\alpha/2} + z_\gamma \sqrt{B} \right)^2 / A. \quad (5)$$

Note that these expressions are heavily dependent on r , the proportion of the risk period to the total observation period.

For the CS model with J age groups, the following test statistic has a similar form [11]:

$$\tilde{T} = \text{sgn}(\hat{\beta}) \sqrt{2} \left[x\hat{\beta} - \sum_{j=1}^J m_j \log(r_j \hat{\rho} + 1 - r_j) \right]^{1/2},$$

where $r_j = e^{\delta_j} e_1 / \sum_{s=1}^J e^{\delta_s} e_s$ and m_j is the total number of events in individuals that are exposed in the j^{th} age group. The sample size n is given by (5) with A and B replaced by

$$\tilde{A} = 2 \sum_{j=1}^J \nu_j [\omega_j \beta - \log(r_j \rho + 1 - r_j)] \quad \text{and} \quad \tilde{B} = \left(\beta^2 / \tilde{A} \right) \sum_{j=1}^J \nu_j \omega_j (1 - \omega_j), \quad (6)$$

where $\nu_j = p_j (r_j \rho + 1 - r_j) / \left[p_0 + \sum_{s=1}^J p_s (r_s \rho + 1 - r_s) \right]$ is the probability a case is exposed in the j^{th} age group and $\omega_j = r_j \rho / (r_j \rho + 1 - r_j)$ is the probability of an event occurring in the risk period given an individual is exposed in the j^{th} age group.

Note that the sample size n refers to the number of events. To obtain the number of cases/subjects, say n_c , an estimate of the cumulative incidence over the observation period, denoted Λ , is needed. Then $n_c = n\Lambda^{-1}(1 - \exp(-\Lambda))$. However, in most applications $n_c \approx n$ since Λ is small [11].

3 Sample size determination when timing of exposure is measured with error

Methods to correct for bias in the timing of exposure onset in the CS model [9] and to provide consistent estimation of the effect sizes is important since, for example, severe underestimation of the magnitude of the true effects can be misleading. However, bias-corrected estimators are more variable than the biased (uncorrected/naive) estimator. When the test of the hypothesis of no effect (i.e., $\beta = 0$) based on the naive estimate is valid, it is typically more powerful than a test based on the bias-corrected estimate. This phenomenon is due to the bias-variance tradeoff well-known in measurement error problems [10]. Thus we focus on sample size calculation methods for the measurement error case series models as presented in Section 2.1 (with and without age effects) using the naive estimate of β , i.e., the standard (uncorrected) CS estimate ignoring exposure onset measurement error. We note that a valid test (asymptotically) is one in which its Type I error rate approaches its nominal test level. Thus, we illustrate the validity of the naive test in both MECS models with and without age effects.

3.1 Method for MECS model without age effects

We first consider sample size calculations for the measurement error case series model when age effects are not included in the model. The objective is to determine the number of events needed to achieve 100 γ percent power for an average amount of exposure onset measurement error, μ_u , at test level α . At the planning stage, specification of μ_u is needed in addition to the standard quantities: the effect size, β , total observation follow-up length, risk period length, e_1 , and proportion of individuals in the population exposed, p . It is useful to define the *relative measurement error* (RME) as μ_u/e_1 ; for instance, in the infection-cardiovascular example introduced earlier, an estimate of μ_u is about 6 days and the risk period length of interest is 30 days after an infection, so $100 \times \text{RME} = 20\%$. We also note that in many applications, p is moderate to high (e.g., the proportion of children who receive a vaccination by age 2 in the U.S. or the proportion of incident dialysis patients age 65 or older who have an infection-related hospitalization).

The naive estimator $\hat{\beta}^*$ is obtained by applying the case series method to data with exposure onset measurement error (while ignoring the error). Because it is a naive MLE, it is asymptotically normal. Mohammed *et al.* [9] characterized the bias of the naive estimator $\hat{\beta}^*$ under the MECS model more generally, and for the current model, $\hat{\beta}^*$ targets β^* , which is an attenuation of the true log relative incidence β (details below). The naive test refers to testing $H_0 : \beta = 0$ without correcting for exposure onset measurement error. Under exposure onset measurement error, the test statistic (T presented in Section 2.2) is also asymptotically distributed as standard normal under the null and under the alternative hypothesis, $T \approx N\left(\text{sgn}(\beta^*) \sqrt{nA^*}, B^*\right)$, where A^* and B^* are obtained from A and B in (4) with β replaced by β^* given in (9) below. That is,

$$A^* = \frac{2p\tilde{\rho}^*}{1+pr(\rho^*-1)} \left[\frac{\rho^*r\beta^*}{\tilde{\rho}^*} - \log(\tilde{\rho}^*) \right] \quad \text{and} \quad B^* = \frac{\beta^{*2}}{A^*} \frac{p}{1+pr(\rho^*-1)} \frac{\rho^*r(1-r)}{\tilde{\rho}^*}, \quad (7)$$

where $\tilde{\rho}^* = \rho^*r + 1 - r$ and $\rho^* = \exp(\beta^*)$ is the relative incidence that the naive CS estimator targets.

Consequently, the sample size needed to achieve 100 γ percent power at level α is

$$n = \left(z_{1-\alpha/2} + z_\gamma \sqrt{B^*} \right)^2 / A^*. \quad (8)$$

It has been shown by Mohammed *et al.* [9] that the naive CS estimator $\hat{\beta}^*$ targets

$$\beta^* = \log \left\{ \frac{e_1 e^\beta + \mu_u (1 - e^\beta)}{e_0 - \mu_u (1 - e^\beta)} \right\} - \log \left(\frac{r}{1-r} \right).$$

Note that β^* depends on the mean of the distribution of u , namely μ_u , and not on the actual distribution of the measurement error under the MECS model. Therefore, at the planning stage, the only additional parameter needed for designing case series studies with exposure onset measurement error is μ_u .

For the proposed sample size calculation method (8), based on the naive test, to be valid, the Type I error rate of the naive test should approach its nominal test level α . This holds for the MECS model without age effects, as it can be seen from (9) that $\beta = 0$ implies that $\beta^* = 0$. Results from simulation studies illustrating that the observed nominal test level is achieved are discussed in Section 4.2.

Naturally, the naive test will have some loss in power compared to the corresponding test based on optimal data where the exposure times are measured precisely (which is not available). This loss of power is examined in the simulation studies of Section 4, where we also evaluate the accuracy of the proposed sample size formula (8).

3.2 Method for MECS model with age effects

We now generalize the sample size calculation method above to measurement error case series models with age effects. This is important since the incidence of events may vary with age in some studies and, hence, adjustment for age effects may be of interest. The simulation studies in Musonda *et al.* [11] suggest that age effects can have a large impact on sample size calculation in case series models. This generalization will require obtaining the targets

of the naive MLEs, namely $\hat{\beta}^*$ and $\hat{\delta}^* = (\hat{\delta}_2^*, \dots, \hat{\delta}_J^*)$, for the log relative incidence in the risk period and the age-specific relative incidences (relative to age group 1), respectively.

That is, the MLE $(\hat{\beta}^*, \hat{\delta}^*)$ is consistent for (β^*, δ^*) , which satisfies the estimating equations (3) in expectation, i.e., (β^*, δ^*) is a solution to the set of J equations:

$$\begin{aligned} \sum_{i=1}^N \sum_{j=1}^J \left[E(\tilde{n}_{ij1}) - n_{i..} \pi_{ij1}^* \right] &= 0, \\ \sum_{i=1}^N \sum_{k=0}^1 \left[E(\tilde{n}_{ijk}) - n_{i..} \pi_{ijk}^* \right] &= 0, \quad j=2, 3, \dots, J, \end{aligned} \quad (10)$$

with $\pi_{ijk}^* = e_{ijk} \exp(\delta_j^* + \beta^* k) / \sum_{s=1}^J \sum_{t=0}^1 e_{ist} \exp(\delta_s^* + \beta^* t)$. For MECS models with age effects, it can be shown that Theorem 1 in Mohammed *et al.* [9] can be extended to obtain the following expectations $E(\tilde{n}_{ijk})$ appearing in (10):

$$\begin{aligned} E(\tilde{n}_{ij0}) &= n_{i..} \frac{e_{ij0} e^{\delta_j} + L_{ij} \mu_u (e^{\delta_j + \beta} - e^{\delta_j})}{\sum_{s=1}^J \sum_{t=0}^1 e_{ist} e^{\delta_s + \beta t}}, \\ E(\tilde{n}_{ij1}) &= n_{i..} \frac{e_{ij1} e^{\delta_j + \beta} + L_{ij} \mu_u (e^{\delta_j} - e^{\delta_j + \beta})}{\sum_{s=1}^J \sum_{t=0}^1 e_{ist} e^{\delta_s + \beta t}}, \quad j=1, 2, \dots, J, \end{aligned}$$

where L_{ij} is the total number of exposures in the j^{th} age group for the i^{th} individual. The set of equations (10) can be solved numerically to obtain (β^*, δ^*) , the true targets of the naive estimators. The Appendix section provides a Newton-Raphson method to solve (10). (Alternatively, note that one can obtain (β^*, δ^*) via simulations.) Once (β^*, δ^*) are computed, they can be substituted into the expressions for $\tilde{A}$ and $\tilde{B}$ given in (6); thus, we define

$$\tilde{A}^* = 2 \sum_{j=1}^J \nu_j^* \left[\omega_j^* \beta^* - \log(r_j^* \rho + 1 - r_j^*) \right], \quad \text{and} \quad \tilde{B}^* = (\beta^{*2} / \tilde{A}) \sum_{j=1}^J \nu_j^* \omega_j^* (1 - \omega_j^*). \quad (11)$$

The sample size formula (8) given above for the MECS model without age effects can be used when the model includes age groups by replacing A^* and B^* with $\tilde{A}^*$ and $\tilde{B}^*$ given in (11).

For MECS models with age effects, as with the model without age effects, the sample size formula rests on the validity of the naive test. Details of simulation studies illustrating the observed nominal test level with different age effect patterns targeting test level are presented in Section 4.2.

4 Simulation studies: Accuracy of sample size formulas and assessment of power

The simulation studies and results summarized in this section serve two main purposes. First, an extensive set of studies are used to assess the accuracy of the sample size formulas for the MECS models (both with and without age effects). Second, several studies were designed to illustrate the magnitude of the loss of power as the average amount of measurement error (μ_u) increases, compared to power under the ideal (optimal) situation where exposure onset times are known precisely for all individuals.

4.1 Simulation design

For modeling exposure onset measurement error without age effects, we consider an observation period of 600 days and several risk period lengths: $e_1 = 30, 60$ and 90 days; i.e., $r = e_1/600 = 0.05, 0.1$ and 0.15 , respectively. All individuals have one exposure, i.e., $p = 1$. To accommodate potential relative incidences in practice, we consider a wide range for the true relative incidence, ρ . More specifically, we take ρ to be $0.5, 1.5, 2, 3$ and 5 . The average exposure onset measurement error, μ_u , expressed relative to e_1 , is the relative measurement error (RME), μ_u/e_1 . We take RME to be $10\%, 20\%$ and 30% . We assess the accuracy of the sample size formulas for the conventional study design power of 80% and 90% (i.e., $\gamma = 0.8$ and 0.9) at level $\alpha = 0.05$. We use the sample size formula (8) and round up to the nearest integer. Calculation of the empirical power, the proportion of tests that correctly reject the null hypothesis, is based on 2000 simulated data sets for each of the 90 combinations of parameter settings (r, ρ , RME and γ).

For MECS models with age effects, we partition the observation period into 3, 4 or 5 age groups with length 200, 150 or 120 days, respectively. Also we consider decreasing, symmetric and increasing patterns of age-specific relative incidence for each model. Table 1 summarizes the age effect parameters (δ) for these nine MECS model with age effects (number of age groups by age-specific effect patterns). For these models, we also consider both models with equal and unequal proportions of individuals in the population exposed in the j^{th} age group (p_j), i.e., $p_j = 1/J, j = 1, \dots, J$, or p_j is varying (three cases). The combination of ρ, r and RME are the same as above for the models without age effects; however, we only include the cases with $r = 0.05$ and 0.10 when simulating data with 5 age groups. Thus, a total of 9 models and 180 parameter combinations were examined (r, ρ , RME, number of age groups J and p_j pattern). As above, empirical power is computed based on 2000 data sets for each model-parameter combination.

4.2 Accuracy of sample size formulas

Results for assessing the accuracy of the sample size formula without age effects (8) at 80% and 90% power are presented in Table 2. Provided are the empirical powers based on 2000 data sets, each with the sample size n (the number of events) determined from (8) at $\gamma = 0.8$ and 0.9 . For all parameter combinations, the empirical power is close to the targets 80% and 90% . For example, with 30% relative measurement error, effect size $\rho = 3$, and $r = 0.05, 0.10$ and 0.15 , the corresponding sample sizes given by formula (8) of $n = 206, 127$ and 105 achieve $91.2\%, 91.0\%$ and 88.9% power, respectively (close to the target of 90% ; bolded in Table 2). Note that, as expected, for a fixed amount of RME and r , the sample size decreases as ρ increases; similarly, for a fixed amount of RME and ρ , sample size decreases as r increases.

For the MECS model with age effects, we present the results for 3 age groups to illustrate the level of accuracy of the sample size formula and defer the results for 4 and 5 age groups to the supplemental materials. (See Section 4.4 below.) Recall from Section 3.2 that the sample size formula is simply given by (8) with $(\tilde{A}, \tilde{B})$ replaced by $(\tilde{A}^*, \tilde{B}^*)$ for models with age-specific relative incidences. Table 3 displays the empirical power corresponding to 80% and 90% power when the proportions of individuals in the population exposed are equal

across age groups (i.e., $p_j = 1/3$, $j = 1, 2, 3$). There is good accuracy, similar to the model without age effects; also, the sample size n is similar to the case of no age effects. For example, with 10% RME and $r = 0.15$, sample sizes required to achieve 80% power for $\rho = 1.5, 2$ and 3 are $n = 424, 131$ and 46 , respectively (for the increasing pattern of age effects; bolded in Table 3). (The corresponding sample sizes for the models without age effects are $n = 407, 126$ and 44 from Table 2; bolded in Table 2.) Note that since $p_j = 1/J$, the sample sizes are similar for each pattern of age effects, as expected.

The accuracy when the p_j 's are not equal is similar. For instance, $p_j = (0.1, 0.2, 0.7)$ corresponds to individuals in a population with increasing proportions of exposure (e.g., infection) with increasing age. Empirical powers for 80% and 90% when $p_j = (0.1, 0.2, 0.7)$ can be found in Table 4. We note that, as expected for the case where p_j 's are unequal, the pattern of sample size requirement will depend on the true underlying age-specific relative incidences (i.e., the δ_j 's). For example, when the RME is 30%, $r = 0.1$ and $\rho = 1.5$, sample sizes needed to achieve 90% power for increasing, symmetric and decreasing age effects are 1182, 1532 and 1883, respectively (bolded in Table 4). This is in contrast to the above case when p_j 's are all equal to $1/3$ and the sample sizes are about constant ($n = 1341, 1327$, and 1356 ; bolded in Table 3) regardless of pattern of age-specific relative incidence.

Also, as noted in Section 3.2, the naive test is valid; i.e., its Type 1 error achieves the nominal level. Under the null hypothesis ($\beta = 0$) without age effects, Table 5(A) confirms that the empirical level achieves the nominal test level of $\alpha = 0.05$ (5%). This similarly holds for all MECS models with varying patterns of age effects (i.e., increasing, symmetric and decreasing age effect patterns) and with equal and unequal p_j 's; see Table 5(B,C). Finally, we note that the results presented here are for the case where the distribution of the exposure onset measurement error u is uniformly distributed with parameters $(\mu_u - 2, \mu_u + 2)$ where μ_u depends on r and RME, as described in the simulation design of Section 4.1. For example, for $r = 0.05$ and 20% RME, $\mu_u = 6$ and $u \sim \text{Uniform}(4, 8)$. Results when u is not uniformly distributed are similar and addressed in Section 4.4.

4.3 Loss of power compared to the optimal case of no exposure onset measurement error

We next illustrate the expected relative loss of power due to exposure onset measurement error. Figure 1 plots the power when RME = 10%, 20% and 30% along with the ideal case of no exposure onset measurement error (RME = 0) for $\rho = 3$ and $r = 0.05$ for MECS model (A) without and (B) with age effects. (Details can be found in Table 6.) Not surprisingly, there is a significant cost associated in using data with measurement error. For example (in the model without age effects), $n = 100$ attains 90.5% power and with 10% RME, the power reduces to 84%, although it is still above the conventional minimum required design power of 80%. However, the power drops to below acceptable levels of 77.1% and 66.5% when the RME is at 20% and 30%, respectively. The sample size n needs to double to about 200 to achieve ~ 90% power, the level for data where exposure onset is measured precisely.

As expected, in order for data with exposure onset measurement error to achieve a similar power compared to the data with precisely known dates/times of exposures, the sample size must be increased. Without exposure onset measurement error (and model without age effects), to achieve a power of 80% when $r = 0.1$, the sample size required is 443, 134, and

46 for $\rho = 1.5, 2$ and 3 , respectively. When the RME is at 10%, the sample sizes required are 553, 166 and 56, respectively, representing an increase of $\sim 22 - 25\%$ to achieve the same power as data without error. In a more extreme case, when the RME is at 30%, the required sample sizes to achieve 80% power increase to 955, 283 and 94, respectively. This is more than twice the sample sizes required when the exposure onset times are measured precisely.

4.4 Supplemental results

As mentioned earlier, additional results are provided in the supplemental materials available at http://dnguyen.ucdavis.edu/html/mecs_design_sup.pdf. This includes results for 4 and 5 age groups, for different patterns of p_j and for assessing the accuracy of the proposed sample size method when the distribution of the exposure onset measurement error u is not uniform. The supplemental results summarize the accuracy for u distributed as normal and gamma; accuracy of the sample size formulas is robust to the distribution of u .

5 Examples and power/sample size calculation tools

Infection and cardiovascular events (e.g., myocardial infarction, unstable angina, stroke or transient ischemic attack) are common in patients with end-stage renal disease on dialysis. It is of interest to determine if the risk of cardiovascular events is increased within a specified window of time after an infection. The case series model can be applied to data from the United States Renal Data System (USRDS) [6] to study this relationship or other longitudinal observational databases more generally (including other medical claims database and electronic medical records systems). The USRDS is an ongoing national longitudinal research database, containing data on nearly all ($> 99\%$) patients with end-stage renal disease in the United States that is utilized by a large community of researchers, especially in nephrology. The database consists of information on hospitalizations including admission and discharge dates and discharge diagnoses. For example, if a hospitalization had an infection discharge diagnosis, then we can reasonably assume that an infection likely occurred during that hospitalization. Under this assumption, we use the time/date of discharge as a marker of the time of infection [5, 9]; thus, the time of infection onset is not known exactly.

To determine power or sample size at the planning stage, it is necessary to have an estimate of the average amount of exposure onset measurement error, μ_u , from subject-specific knowledge and/or preliminary data (similar to other measurement error problems). For this example, we can reasonably estimate μ_u from the length of hospital stay (the difference between the hospitalization discharge and admission times). More precisely, assuming equal likelihood of infection during a hospitalization stay, $u_i|l_i \sim \text{Uniform}(0, l_i)$, where l_i is the duration/length of hospitalization for subject i . Since $\mu_u = E(u) = E\{E(u_i|l_i)\}$, a consistent estimate is $0.5 \sum_i^N l_i / N \sim 11/2 = 5.5$ days. Thus, we illustrate the sample size formulas for $\mu_u = 4$ and 8 , which represent optimistic (low) and high amounts of measurement error in this application.

Suppose that our observation period is 300 days, the risk period length of interest is $e_1 = 30$ days (after an infection) and we assume three age groups with lengths $e_j = 100$ days. We set

the proportions of exposure in each age group to $p_1 = 0.2$, $p_2 = 0.3$ and $p_3 = 0.5$. The age-specific relative incidences are $\exp(\delta_1) = 1$, $\exp(\delta_2) = 1.5$ and $\exp(\delta_3) = 2$. We wish to be able to detect a true relative incidence of cardiovascular events in the 30-day risk period of $\rho = 1.5$ at $\alpha = 0.05$ with 90% power. Using $z_{0.9} = 1.2816$ and $z_{0.975} = 1.96$,

$$n = \left(z_{0.975} + z_{0.9} \sqrt{\tilde{B}^*} \right)^2 / \tilde{A}^*$$
, where $(\tilde{A}^*, \tilde{B}^*)$ are given by (11). With $\mu_u = 4$ and 8 (i.e., RME 13.3 and 26.7%), $(\tilde{A}^*, \tilde{B}^*) = (0.01425, 1.08447)$ and $(0.00971, 1.07148)$, respectively; this gives $n = 762$ and $n = 1113$. Similarly, without age effects, $(A^*, B^*) = (0.01345, 1.08960)$ and $(0.00932, 1.07634)$ and the required sample sizes are $n = 809$ and $n = 1162$ for $\mu_u = 4$ and 8.

The proposed sample size (or power) calculation method for the measurement error case series models as well as the original case series models for precisely measured exposure onset is implemented in JavaScript and is made publicly available. Also, R functions to include age effects for the MECS models, as described in Section 3, are made freely available for public use. These tools can be accessed at http://dnguyen.ucdavis.edu/html/MECS_Design/index.html.

6 Discussion

In this work we proposed and evaluated a method for sample size determination for the measurement error case series models [9], where the (timing of) exposure onset is measured with error. The approach is based on the naive test, where one applies the standard case series method to the data measured with error to obtain the naive MLE. We also illustrate the relative loss of power attributed to exposure onset measurement error, compared to the ideal situation where exposure times are measured perfectly.

Several model simplifications and assumptions were made paralleling the work of [11], which appears reasonable at the design stage. One of these is the assumption of equal observation periods. However, we also examined simulation studies with unequal observation periods as well and the results similarly hold. If one can obtain an estimate of the average length of observation period from preliminary data or posit an expected/average observation length across subjects, then this can be used in the proposed sample size calculation and our simulations show similar empirical power (results not shown). Secondly, we have assumed only one risk period and in the models with age effects, that the true and observed risk periods are fully contained within one age group. This method would need to be modified to include cases with multiple risk periods or where risk periods can overlap multiple age groups. However, in practice, it is reasonable to consider the simpler case and focus on one main risk period of interest that is shorter than the age group lengths at the design stage. It is also reasonable at the design stage to consider simplified models with a discrete exposure [11] as we did in this work, although we note that a more general case series model for continuous exposures exists [12]. Finally, publicly available tools for the proposed method were implemented to facilitate case series studies and the tools can be used to easily examine the sensitivity of the effect of exposure onset measurement error on study power.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank two reviewers and an associate editor for their suggestions to improve the paper. This publication was made possible by grant UL1 RR024146 from the National Center for Advancing Translational Sciences (DVN, LSD) and partially by NIDDK grant DK092232 (DS, LSD, DVN). We thank Yi Mu and Vien Nguyen in the UC Davis Department of Public Health Sciences. The interpretation and reporting of the data presented here are the responsibility of the authors and in no way should be seen as an official policy or interpretation of the United States government. This study was approved by the Institutional Review Board of the University of California, Davis Health System.

Appendix A

Calculating the targets of the naive CS estimators

The proposed power/sample size determination described in Section 3 requires computation of (A^*, B^*) in (7) and $(\tilde{A}^*, \tilde{B}^*)$ in (11), which requires computation of β^* and (β^*, δ^*) , respectively, for the MECS models without and with age effects. There are two ways to do this. One can directly use Monte Carlo simulation to generate data from MECS models and apply the standard CS MLE. Here one option is to use averages of the estimates of (β^*, δ^*) for large n . Alternatively, one can directly obtain (β^*, δ^*) without simulation by solving the set of J equations given in (10) of Section 3.2. Recall $\pi_{ijk}^* = e_{ijk} e_{j^*}^{\delta_j^* + \beta^* k} / \sum_{s,t} e_{ist} e_c^* e_{j^*}^{\delta_j^* + \beta^* t}$ and let

$$\begin{aligned} a_1 &\equiv \sum_{i=1}^N \sum_{j=1}^J \left[E(\tilde{n}_{ij1}) - n_{i..} \pi_{ij1}^* \right] = 0, \\ b_1 &\equiv \sum_{i=1}^N \sum_{k=0}^1 \left[E(\tilde{n}_{ijk}) - n_{i..} \pi_{ijk}^* \right] = 0, \quad j=2, 3, \dots, J, \end{aligned}$$

where $E(\tilde{n}_{ij0})$ and $E(\tilde{n}_{ij1})$ were given in Section 2.1. This set of equations can be solved numerically for (β^*, δ^*) by the Newton-Raphson method where the update of (β^*, δ^*) at

iteration $t+1$ is $(\beta^*, \delta^*)^{(t+1)} = (\beta^t) - (\mathbf{J}^{(t)})^{-1} \mathbf{d}^{(t)}$, with $\mathbf{d}^{(t)} = (a_1^{(t)}, b_2^{(t)}, \dots, b_J^{(t)})^T$ and $\mathbf{J}^{(t)}$ is a $J \times J$ matrix of partial derivatives evaluated at $(\delta^*, \delta^*)^{(t)}$:

$$\begin{aligned} \frac{\partial a_1}{\partial \beta^*} &= - \sum_{i=1}^N n_{i..} \pi_{i.1}^* (1 - \pi_{i.1}^*), \\ \frac{\partial b_j}{\partial \delta_j^*} &= - \sum_{i=1}^N n_{i..} \pi_{ij.}^* (1 - \pi_{ij.}^*), \quad j=2, \dots, J \\ \frac{\partial b_j}{\partial \delta_l^*} &= - \sum_{i=1}^N n_{i..} \pi_{ij.}^* \pi_{il.}^*, \quad j \neq l \\ \frac{\partial a_1}{\partial \delta_j^*} &= \frac{\partial b_j}{\partial \beta^*} = - \sum_{i=1}^N n_{i..} (\pi_{ij1}^* - \pi_{ij.}^* \pi_{ij.}^*), \quad j=2, \dots, J. \end{aligned}$$

An R function that makes the above computations is provided as part of the sample size calculation function, available at [http://dnguyen.ucdavis.edu/.html/MECS Design/index.html](http://dnguyen.ucdavis.edu/.html/MECS%20Design/index.html).

References

1. Farrington CP. Relative incidence estimation from case series for vaccine safety evaluation. *Biometrics*. 1995; 51:228–235. [PubMed: 7766778]
2. Farrington CP, Nash J, Miller E. Case series analysis of adverse reactions to vaccines: a comparative evaluation. *American Journal of Epidemiology*. 1995; 143:1165–1173. (Erratum 1998; 147:93). [PubMed: 8633607]
3. Gibson JE, Hubbard RB, Smith CJP, Tata LJ, Britton JR, Fogarty AW. Use of self-controlled analytical techniques to assess the association between use of prescription medications and the risk of motor vehicle crashes. *American Journal of Epidemiology*. 2009; 169:761–768. [PubMed: 19181876]
4. Smeeth L, Thomas SL, Hall AJ, Hubbard R, Farrington P, Vallance P. Risk of myocardial infarction and stroke after acute infection or vaccination. *New England Journal of Medicine*. 2004; 351:2611–2618. [PubMed: 15602021]
5. Dalrymple LS, Mohammed SM, Mu Y, Johansen KL, Chertow GM, Grimes B, Kaysen GA, Nguyen DV. The risk of cardiovascular-related events following infection-related hospitalizations in older patients on dialysis. *Clinical Journal of the American Society of Nephrology*. 2011; 6:1708–1713. [PubMed: 21566109]
6. US Renal Data System. *USRDS 2010 Annual Data Report: Atlas of Chronic Kidney Disease and End-Stage Renal Disease in the United States*. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases; Bethesda, MD: 2010.
7. Whitaker HJ, Hocine M, Farrington CP. The methodology of self-controlled case series studies. *Statistical Methods in Medical Research*. 2009; 18:7–26. [PubMed: 18562396]
8. Whitaker HJ, Farrington CP, Spiessens B, Musonda P. Tutorial in biostatistics: The self-controlled case series method. *Statistics in Medicine*. 2006; 25:1768–1797. [PubMed: 16220518]
9. Mohammed SM, Senturk D, Dalrymple DS, Nguyen DV. Measurement error case series models with application to infection-cardiovascular risk in older patients on dialysis. *Journal of the American Statistical Association*. 2012 in-press.
10. Carroll, RJ.; Ruppert, D.; Stefanski, LA.; Crainiceanu, CM. *Measurement Error in Nonlinear Models: A Modern Perspective*. Chapman and Hall/CRC: Boca Raton; 2006.
11. Musonda P, Farrington CP, Whitaker HJ. Sample sizes for self-controlled case series studies. *Statistics in Medicine*. 2006; 25:2618–2631. (Erratum 2008; 27:4854–4855). [PubMed: 16372391]
12. Farrington CP, Whitaker HJ. Semiparametric analysis of case series data. *Applied Statistics*. 2006; 55:553–594.

Figure 1.

Power curves illustrating the relative loss in power with increased relative measurement error (RME) in data without (A) and with (B) age groups (with increasing age effects) when $r = 0.05$ and $e^{\beta} = 3$.

Table 1

Nine measurement error case series models with increasing, symmetric and decreasing age effect (age-specific relative incidence) patterns used in the simulation studies.

Number of age groups (J)	Effect pattern	Age-specific relative incidence				
		$e^{\delta 1}$	$e^{\delta 2}$	$e^{\delta 3}$	$e^{\delta 4}$	$e^{\delta 5}$
3	Increasing	1	2	3	-	-
	Symmetric	1	2	1	-	-
	Decreasing	1	1/2	1/3	-	-
4	Increasing	1	2	3	4	-
	Symmetric	1	2	2	1	-
	Decreasing	1	1/2	1/3	1/4	-
5	Increasing	1	2	3	4	5
	Symmetric	1	2	3	2	1
	Decreasing	1	1/2	1/3	1/4	1/5

Table 2

Empirical power based on 2000 simulated data sets for MECS models without age effects for varying amounts of relative measurement error (RME).

RME	<i>r</i>	ρ	80%		90%	
			<i>n</i>	Power	<i>n</i>	Power
0.1	0.05	0.5	630	83.5	828	90.8
		1.5	1006	80.3	1360	89.8
		2	295	81.3	401	89.2
		3	95	78.7	130	90.4
		5	35	81.3	48	91.4
	0.1	0.5	326	82.3	430	91.4
		1.5	553	81.6	747	89.3
		2	166	81.3	226	91.7
		3	56	79.9	77	89.6
		5	22	77.2	30	90.3
	0.15	0.5	226	77.0	298	92.7
		1.5	407	79.8	549	89.9
		2	126	77.1	170	87.5
		3	44	84.1	60	92.4
		5	19	78.9	25	91.7
0.2	0.05	0.5	835	85.3	1100	91.9
		1.5	1272	78.9	1718	89.2
		2	369	78.1	502	88.6
		3	118	76.3	161	86.9
		5	43	78.0	59	91.7
	0.1	0.5	439	83.7	580	91.1
		1.5	712	79.3	961	89.6
		2	213	78.2	288	90.1
		3	71	80.3	97	92.1
		5	28	78.0	38	92.8
	0.15	0.5	310	81.2	410	90.2
		1.5	535	81.3	721	89.5
		2	164	83.5	221	92.0
		3	57	76.4	78	89.6
		5	24	80.2	32	87.9
0.3	0.05	0.5	1145	85.3	1513	91.8
		1.5	1667	78.2	2249	90.0
		2	479	79.8	650	91.0
		3	151	78.2	206	91.2
		5	54	77.5	74	90.9
	0.1	0.5	615	82.0	814	92.0
		1.5	955	77.4	1287	89.3

RME	r	ρ	80%		90%	
			n	Power	n	Power
0.15		2	283	79.9	382	89.4
		3	94	81.4	127	91.0
		5	36	79.2	49	89.9
	0.5	0.5	445	82.5	590	90.4
		1.5	738	79.1	993	89.2
		2	224	78.5	303	89.2
		3	78	78.1	105	88.9
		5	32	75.3	44	91.3

Table 3

Empirical power corresponding to (A) 80% and (B) 90% for models with 3 age groups, varying amounts of relative measurement error (RME) and $\rho_j = (1/3, 1/3, 1/3)$.

RME	r	ρ	(A) Incr.			Symm.			Decr.			(B) Incr.			Symm.			Decr.		
			n	Power	n	Power	n	Power	n	Power	n	Power	n	Power	n	Power	n	Power	n	Power
0.1	0.05	0.5	636	84.4	634	84.2	638	83.0	836	91.3	834	92.0	838	92.6						
			1.5	1018	78.4	1015	81.3	1021	80.3	1375	90.4	1371	91.1	1380	90.5					
		2	299	80.4	298	80.5	300	77.4	406	89.9	405	89.3	407	90.8						
		3	96	79.4	96	81.0	97	80.5	132	90.8	132	89.6	132	88.9						
	0.1	5	35	77.9	35	77.1	35	76.4	48	88.0	48	88.8	49	88.9						
			0.5	333	82.0	332	80.6	335	82.9	439	90.2	436	92.8	442	92.1					
		1.5	567	80.5	564	80.2	571	80.1	765	89.9	760	90.3	771	90.2						
		2	171	78.8	170	80.5	172	77.4	232	90.3	230	89.0	233	89.7						
	0.15	3	58	78.8	57	78.4	58	78.0	79	89.2	78	90.8	79	89.7						
			5	23	80.2	23	79.7	23	79.5	31	90.0	31	91.0	31	88.3					
		0.5	234	83.1	232	80.4	237	82.7	309	90.0	306	91.9	312	90.8						
		1.5	424	79.2	420	80.1	429	79.0	571	89.3	565	89.1	578	89.5						
0.2	0.05	2	131	78.6	130	78.0	132	79.1	177	89.3	175	89.4	179	89.9						
			3	46	80.7	46	79.3	47	77.7	62	87.3	62	88.4	63	89.4					
		5	19	78.6	19	79.0	20	78.9	26	89.3	26	90.2	26	89.8						
		0.5	845	82.8	842	84.4	847	81.9	1114	91.8	1110	91.7	1117	91.7						
	0.1	1.5	1291	81.0	1286	78.4	1297	79.2	1743	89.6	1737	89.4	1751	89.5						
			2	375	80.9	374	78.7	377	80.9	510	90.1	508	88.9	512	88.9					
		3	120	80.1	119	78.1	121	79.0	164	89.9	163	89.8	165	89.3						
		5	43	76.2	43	75.7	44	77.7	60	89.2	59	90.4	60	89.5						
	0.15	0.5	451	82.1	448	80.4	454	82.0	595	91.8	592	91.8	600	91.2						
			1.5	735	81.1	729	79.9	742	81.3	992	89.7	983	90.6	1001	89.1					
		2	220	77.7	218	79.9	222	77.5	298	89.9	295	90.0	300	89.2						
		3	74	79.7	73	79.2	75	79.5	100	90.2	99	89.8	101	90.3						
0.15	0.5	5	29	80.3	29	79.8	29	77.4	39	88.8	39	89.1	40	89.7						
		0.5	324	83.2	321	81.4	328	80.3	428	91.1	424	91.5	434	91.0						

RME	r	ρ	(A) Incr.			Symm.			Decr.			(B) Incr.			Symm.			Decr.		
			n	Power	n	Power	n	Power	n	Power	n	Power	n	Power	n	Power	n	Power		
0.3	0.05	1.5	564	79.9	556	79.2	572	79.2	759	89.9	749	90.3	770	89.9						
		2	173	79.9	170	79.4	176	77.8	234	90.9	230	89.7	237	90.0						
		3	61	81.0	60	81.4	62	79.1	82	90.8	81	89.2	83	88.8						
		5	25	78.8	25	80.7	26	78.8	34	88.7	34	90.3	35	90.2						
		0.5	1162	82.9	1157	83.3	1167	82.7	1536	91.3	1530	91.6	1542	91.6						
	0.1	1.5	1699	80.9	1690	79.5	1708	80.2	2291	89.1	2280	88.2	2304	88.8						
		2	489	79.1	487	80.3	492	78.7	663	88.3	660	90.1	667	89.6						
		3	155	81.0	154	80.5	156	79.7	211	89.5	210	90.5	212	88.9						
		5	55	77.7	55	77.5	56	79.1	76	89.6	75	89.1	76	87.6						
		0.5	636	82.8	631	81.5	642	81.1	842	90.8	835	90.8	850	91.4						
	0.15	1.5	995	79.0	985	79.0	1007	81.1	1341	88.9	1327	88.7	1356	89.4						
		2	295	80.6	292	79.5	299	79.2	399	90.9	394	88.8	404	89.3						
		3	98	80.1	97	79.6	99	78.7	133	90.3	131	88.1	135	89.7						
		5	38	78.6	38	79.8	39	78.9	52	90.0	51	90.6	53	89.3						
		0.5	471	80.6	465	81.0	480	81.1	625	90.4	616	90.2	636	90.1						
		1.5	789	80.6	775	78.8	804	79.7	1062	90.6	1043	89.4	1082	90.5						
		2	240	81.3	236	78.9	245	80.6	324	90.9	318	89.9	330	89.6						
		3	84	79.4	82	79.0	85	78.0	113	88.7	111	89.5	115	88.6						
		5	35	79.0	34	81.1	35	79.6	47	89.5	46	89.6	48	88.4						

Table 4

Empirical power corresponding to (A) 80% and (B) 90% for models with 3 age groups, varying amounts of relative measurement error (RME) and $p_j = (0.1, 0.2, 0.7)$.

RME	r	ρ	(A) Incr.			Symm.			Decr.			(B) Incr.			Symm.			Decr.		
			n	Power	n	Power	n	Power	n	Power	n	Power	n	Power	n	Power	n	Power	n	Power
0.1	0.05	0.5	495	81.3	706	83.6	893	83.7	651	91.4	927	92.4	1174	92.0						
			1.5	806	78.9	1124	80.1	1407	80.5	1089	90.1	1519	90.1	1903	90.4					
		2	238	79.8	329	81.0	410	79.0	324	89.8	447	90.3	558	88.9						
			3	78	79.4	106	79.0	131	79.7	107	89.1	145	90.3	179	90.4					
		5	29	72.4	38	76.7	47	77.9	40	85.5	53	89.5	65	89.5						
			0.1	0.5	264	82.1	369	80.6	466	80.1	347	90.6	486	90.2	613	90.7				
	0.15	1.5	465	79.8	623	80.6	770	80.3	627	89.6	841	90.1	1040	90.5						
			2	142	79.1	187	80.5	229	79.4	192	89.2	254	88.3	311	87.4					
		3	49	79.8	63	79.3	76	78.9	67	89.2	86	88.1	104	89.3						
			5	20	74.3	25	79.8	29	78.1	27	88.3	34	88.8	40	89.2					
		0.15	0.5	190	79.8	260	80.7	326	80.7	251	89.9	342	89.4	430	88.9					
			1.5	363	79.0	465	77.6	565	77.9	488	87.9	626	88.3	762	87.4					
0.2	0.05	2	114	79.9	143	79.1	172	79.8	154	88.2	193	90.0	232	88.7						
			3	42	78.8	50	78.1	59	77.3	56	89.2	68	89.2	80	87.0					
		5	18	79.8	21	80.0	24	78.0	24	87.3	28	88.0	32	87.5						
			0.5	665	83.2	942	83.4	1192	82.5	877	90.9	1243	92.3	1571	92.5					
		1.5	1036	79.8	1435	79.5	1797	79.0	1399	89.7	1939	90.4	2427	89.0						
			2	304	80.3	417	78.8	519	79.2	412	90.1	566	89.7	706	89.9					
	0.1	3	99	78.0	133	77.9	164	78.7	135	90.1	182	88.0	225	89.0						
			5	37	76.5	48	79.1	58	78.6	50	87.9	66	88.2	80	88.5					
		0.5	367	81.0	507	81.6	637	80.7	485	90.5	669	90.0	841	91.1						
			1.5	622	80.0	822	79.3	1013	79.7	838	89.6	1108	90.4	1366	89.3					
		2	189	78.7	245	78.5	300	78.6	255	89.8	332	90.9	406	89.2						
			3	65	79.5	82	79.7	99	78.8	88	89.5	112	88.8	135	88.6					
0.15	0.5	5	26	79.7	32	79.3	38	77.5	36	88.4	44	88.7	52	89.0						
			277	78.9	368	80.5	460	80.6	367	88.4	487	90.1	607	90.3						

RME	r	ρ	(A) Incr.			Symm.			Decr.			(B) Incr.			Symm.			Decr.		
			n	Power	n	Power	n	Power	n	Power	n	Power	n	Power	n	Power	n	Power		
0.3	0.05	1.5	511	77.4	637	78.6	772	78.2	687	88.5	858	88.9	1040	88.5						
		2	160	79.7	195	77.6	234	78.1	216	89.0	264	89.1	316	86.8						
		3	58	78.2	68	77.7	80	78.9	78	89.1	93	88.9	109	88.2						
		5	25	77.9	29	78.1	33	78.3	34	89.0	39	87.8	44	86.8						
		0.5	928	83.0	1306	82.3	1649	83.3	1227	91.6	1727	91.6	2180	92.1						
	0.1	1.5	1386	77.6	1905	80.6	2382	78.7	1869	88.6	2571	89.4	3214	88.7						
		2	403	80.3	548	76.8	682	78.0	546	89.3	744	89.9	926	90.2						
		3	129	78.8	173	79.0	213	78.9	176	90.5	236	88.2	292	89.5						
		5	48	77.3	62	78.2	75	78.6	65	89.0	85	87.8	103	88.6						
		0.5	538	79.1	728	82.4	912	80.6	713	90.9	964	91.7	1207	90.8						
	0.15	1.5	878	78.1	1137	77.9	1397	79.1	1182	90.6	1532	89.4	1883	90.2						
		2	265	78.3	337	78.9	411	80.0	357	89.8	456	88.5	556	90.7						
		3	90	78.5	112	78.1	135	77.9	122	89.3	152	89.0	183	89.1						
		5	36	77.6	44	79.0	51	77.0	49	87.6	59	89.1	70	88.5						
		0.5	437	78.3	554	80.4	687	78.7	580	89.6	735	89.0	910	89.8						
		1.5	776	77.4	931	79.8	1120	79.0	1043	88.3	1252	89.5	1508	89.1						
		2	241	78.9	284	78.6	338	77.1	325	88.4	383	88.5	457	87.9						
		3	86	77.9	99	79.0	116	79.0	116	88.6	134	88.3	157	88.5						
		5	37	76.7	42	78.6	47	77.7	50	88.1	56	88.5	64	88.2						

Table 5

Empirical/observed powers (in percent) based on 2000 simulated data sets under the null hypothesis of no effect ($\beta = 0$) for models (A) without age effects and (B,C) with varying patterns of 3 age groups (increasing, symmetric and decreasing) for different amounts of relative measurement error (RME) with (B) $p_j = (1/3, 1/3, 1/3)$ and (C) $p_j = (0.1, 0.2, 0.7)$.

RME	r	(A) No age effects	(B) Age effect pattern			(C) Age effect pattern		
			Incr.	Symm.	Decr.	Incr.	Symm.	Decr.
0.1	0.05	4.9	5.3	4.3	5.2	5.3	5.1	5.1
	0.1	4.9	5.4	5.4	5.7	6.6	4.9	5.5
	0.15	6.0	5.6	5.1	4.5	6.5	6.5	6.2
0.2	0.05	4.2	5.3	4.7	5.5	5.0	5.8	4.9
	0.1	5.8	4.1	5.2	4.7	5.4	4.9	6.6
	0.15	4.5	4.9	5.4	5.2	6.3	5.9	5.8
0.3	0.05	5.1	4.8	5.1	4.8	6.3	6.4	5.5
	0.1	4.5	5.0	5.2	4.8	6.0	5.8	6.0
	0.15	5.6	5.5	4.5	5.2	5.7	6.2	6.7

Table 6

Empirical relative loss of power for models without age effects and with 3 age groups with $e^\delta = (1, 2, 3)$, $r = 0.05$, $e^\beta = 3$, relative measurement error (RME) at 10, 20 and 30%. Optimal power, i.e., power for the case of no exposure onset measurement error (ME) is provided for reference.

p_j	n	RME			
		No ME	10%	20%	30%
No age effects	80	77.0	71.8	62.5	51.7
	100	90.5	84.0	77.1	66.5
	120	90.6	85.7	79.2	66.1
	140	95.7	93.2	84.8	75.2
	160	98.3	96.0	92.0	83.6
	180	98.2	96.4	92.8	84.8
	200	99.3	97.6	94.7	89.8
	300	100.0	99.8	99.3	97.5
(1/3,1/3,1/3)	80	79.5	74.0	61.7	51.6
	100	86.6	81.3	73.4	63.4
	120	90.4	86.2	78.1	70.2
	140	94.8	91.5	85.2	76.0
	160	97.1	93.9	88.9	80.5
	180	98.1	96.1	91.9	85.2
	200	98.8	97.6	93.4	87.7
	300	99.8	99.6	99.1	96.4
(0.1,0.2,0.7)	80	86.8	81.2	71.5	60.5
	100	92.5	87.8	79.4	68.2
	120	95.4	92.1	85.6	75.5
	140	97.7	95.5	90.5	82.0
	160	98.8	97.1	93.7	87.1
	180	99.1	97.9	96.1	89.3
	200	99.6	99.0	97.0	92.3
	300	100.0	99.9	99.8	97.9