Time-to-event estimation of birth year prevalence trends: a method to enable investigating the etiology of childhood disorders including autism

Alexander G. MacInnis[1*]

[1] Independent researcher, Mountain View, California, United States of America

* Corresponding author

Email: a.macinnis@alumni.stanford.edu (AGM)

Abstract

Measures of incidence are essential for investigating etiology. For congenital diseases and disorders of early childhood, birth year cohort prevalence serves the purpose of incidence. There is uncertainty and controversy regarding the birth prevalence trend of childhood disorders such as autism and intellectual disability because changing diagnostic factors can affect the rate and timing of diagnosis and confound the true prevalence trend. The etiology of many developmental disorders is unknown, and it is important to investigate. This paper presents a novel method, Time-to-Event Prevalence Estimation (TTEPE), to accurately estimate the time trend in birth prevalence of childhood disorders correctly adjusted for changing diagnostic factors. There is no known existing method that meets this need. TTEPE is based on established time-to-event (survival) analysis techniques. Input data are rates of initial diagnosis for each birth year cohort by age or, equivalently, diagnostic year. Diagnostic factors form diagnostic pressure, i.e., the probability of diagnosing cases, which is a function of diagnostic year. Changes in diagnostic criteria may also change the effective prevalence at known times. A discrete survival model predicts the rate of initial diagnoses as a function of birth year, diagnostic year, and age. Diagnosable symptoms may develop with age, affecting the age of diagnosis, so TTEPE incorporates eligibility for diagnosis. Parameter estimation forms a non-linear regression using general-purpose optimization software. A simulation study validates the method and shows that it produces accurate estimates of the parameters describing the trends in birth prevalence and diagnostic pressure. The paper states the assumptions underlying the

analysis and explores optional additional analyses and potential deviations from assumptions. TTEPE is a robust method for estimating trends in true case birth prevalence controlled for diagnostic factors and changes in diagnostic criteria under certain specified assumptions.

Keywords

Birth year prevalence; birth year cohort; incidence; time-to-event; survival analysis; diagnostic factors; diagnostic pressure; non-linear regression; autism; intellectual disability; childhood disorders; developmental disorders

Introduction

In epidemiology, incidence - the rate of new cases - is a fundamentally important metric for estimating causal associations of time-varying risk factors with rates of a disorder [1,2]. Incidence is different from prevalence, which is the proportion of a defined population with the disorder at a defined time. For some disorders including congenital diseases and developmental disorders such as autism and intellectual disability, birth year cohort prevalence is used instead of incidence [2,3] because disorder incidence is indistinguishable from birth prevalence, and diagnosis may occur later if at all. Sometimes “incidence” is used to mean the rate of incident diagnoses rather than true incidence, which refers to a disorder. While this usage is understandable because observations inherently represent diagnosis and identification, the difference can be

critically important when there is uncertainty about the rate and timing of diagnosing or identifying cases. Serious developmental disorders such as autism and intellectual disability are important topics of study because they cause a significant reduction in quality of life across the lifespan [4] for the individuals and their families, and they affect large numbers of people.

Studies of time trends in birth year prevalence or incidence are subject to biases resulting from changes in diagnostic factors and diagnostic criteria [2]. Investigators can estimate incidence and birth prevalence directly from data on diagnoses. But concerns that diagnostic factors and diagnostic criteria may have affected the data can lead to a lack of confidence in the validity of direct estimates. Diagnostic factors are those that influence the probability of diagnosing or identifying cases. Examples include awareness, outreach efforts, screening, diagnostic practice, diagnostic criteria, social factors, policies, and financial incentives for diagnosis. Changes in diagnostic criteria can also have the additional effect of changing the effective birth prevalence by including or excluding as cases some portion of the population compared to prior criteria, with the changes occurring when the new criteria take effect.

Consider, for example, this question. If reported birth prevalence increased over time, was this caused by changes in actual birth prevalence, the probability of diagnosing cases, diagnostic criteria affecting prevalence, or some combination of these factors? It

is challenging to disentangle these effects, and there is no known existing method capable of doing so correctly.

Literature Review

There are many studies on the prevalence of developmental disorders, yet very few of them directly address birth year prevalence trends and very few address methods of adjustment for diagnostic factors. The series of reports from the US Centers for Disease Control and Prevention's (CDC) Autism and Developmental Disabilities Monitoring Network (ADDM) [5-13] estimate the prevalence of autism among children who were eight years old at each even-numbered year 2000 through 2016. Each report describes the prevalence of a single year birth cohort, subject to rounding, born eight years before the respective study year. The set of reports represents the trend in birth year prevalence, but the reports describe the findings as simply "prevalence," and do not discuss birth year prevalence or similar names. The ADDM reports suggest that the observed increases in (birth year) prevalence may result from various factors, including changing composition of study sites and geographic coverage, improved awareness, and changes in diagnostic practice and availability of services. However, they do not suggest methods to quantify such effects or to adjust for them. Croen [14] examined birth year prevalence trends in autism and mental retardation in California for birth years 1987 to 1994. They concluded that the data and methods available were insufficient to determine how much of the observed increase reflected an increase in true birth prevalence. Hansen [15] recommends using the cumulative incidence of diagnoses of childhood psychiatric disorders for each 1-year birth cohort as a measure of risk. They

do not suggest an analytical method to estimate or adjust for the effects of diagnostic factors. Nevison [16] presents California Department of Developmental Services data showing a sharp rise in birth year prevalence of autism over several decades but does not discuss methods to adjust for the effects of diagnostic factors.

Elsabbagh [17] states that investigating time trends in prevalence or incidence requires holding diagnostic factors such as case definition and case ascertainment “under strict control over time,” but does not suggest a method for doing so. Campbell [18] reviews prevalence estimates and describes an ongoing controversy about them. They emphasize the distinction between prevalence and incidence but do not mention birth year prevalence. They summarized the CDC ADDM estimates and stated that one cannot infer incidence from the ADDM prevalence estimates. However, they did not mention that the ADDM estimates are birth year prevalence, which serves the purpose of incidence for childhood disorders. Campbell indicates that analyses should control for certain diagnostic factors, but they do not suggest a method for doing so. Baxter [4] examined prevalence and incidence but did not mention birth year prevalence and did not indicate whether their use of “incidence” refers to incident diagnoses or incidence of the disorder. Later sections of this paper show why the distinction is crucial. Baxter adjusted for covariates that they assumed introduced bias, including the use of dichotomous variables representing the most recent diagnostic criteria. Such variables inherently represent the time each set of criteria took effect. However, Schisterman [19] shows that controlling for variables on a causal path from the input (time, in this case) to

the outcome (prevalence or incident diagnoses) constitutes inappropriate adjustment and biases the estimate of the primary effect (i.e., of time on prevalence or incident diagnoses) towards zero. Similarly, Rothman [2] states that controlling for intermediate variables typically causes a bias towards finding no effect.

Keyes [20] used age-period-cohort analysis to attempt to disentangle the effects of birth year (cohort) from diagnostic year (period) and concluded that period effects best explain observed California data. They also argued that period effects represent diagnostic factors, without noting that birth year prevalence is inherently a cohort effect. Spiers [21], in a letter regarding Keyes, pointed out that the method used is extremely sensitive to the constraints specified and could as easily have concluded that period, i.e., diagnostic year, effects best explain the data. Spiers also disputed Keyes' interpretation of cohort effects. King [22] implicitly used an age-period-cohort analysis, assuming that period effects are dominant and controlling for birth year. There is extensive literature on the problems with using age-period-cohort analysis to separate the effects of birth year (cohort) from diagnostic year (period). Rodgers [23] states that a constraint of the type used in Keyes "in fact [it] is exquisitely precise and has effects that are multiplied so that even a slight inconsistency between the constraint and reality, or small measurement errors, can have very large effects on estimates." O'Brien [24], in a book devoted to this topic, states, regarding the relationships of age, period and cohort to the dependent variable, "There is no way to decide except by making an assumption about the relationship between these three variables." MacInnis [25] showed that

diagnostic factors are represented by the years of first diagnoses, formulates the problem as one of separating birth year from diagnostic year, and shows that age-period-cohort approaches are not suitable for such analyses. In particular, implicit assumptions to make the model estimable cause the resulting estimates to conform to the assumptions, forming circular logic.

Campbell [18] and McKenzie [26] both point out that various factors could potentially affect the rate of diagnoses without affecting the true case rate.

Overview

The primary aim of this work is to develop and specify a method to estimate birth prevalence trends, correctly adjusted for trends in the set of diagnostic factors and changes in diagnostic criteria. Armed with such a tool, researchers can quantify the effects of the set of variable causal factors separately from those of the set of diagnostic factors. Where covariates are available, investigators can estimate associations of birth prevalence with a variety of population characteristics that may be causal or explanatory.

This paper presents a novel statistical method called time-to-event prevalence estimation (TTEPE). It uses time-to-event survival analysis to estimate the trend in true birth year prevalence, correctly adjusted for changes in the set of diagnostic factors and diagnostic criteria. It presents the derivation of the analytical method from first principles and states all the underlying assumptions. A simulation study shows that the method

effectively separates and quantifies the birth year trend and the trend in the effects of diagnostic factors, producing accurate estimates.

Method of time-to-event prevalence estimation (TTEPE)

Background

Comparison of prevalence estimates across multiple studies generally is not suitable for informing tends in birth prevalence nor incidence [2]. Different prevalence estimates may use different mixes of birth years and ages, as well as numerous other possible differences between prevalence studies [27]. Many combinations of trends in birth prevalence and diagnostic factors could potentially explain observed prevalence trends. The Introduction section briefly describes the problems with age-period-cohort analyses. Age (A), diagnostic year (DY), and birth year (BY) are exactly collinear, $DY = BY + A$, subject to rounding, which leads to unidentified estimates when using a linear predictor. Another reason is that the age distribution of diagnoses can differ for different solutions of BY and DY, as shown below. It is challenging to estimate the age distribution correctly given the collinearity problem.

Analysis of the cumulative incidence, to a consistent age, of diagnoses in each birth cohort comes closer to estimating the trend in true birth prevalence, but results are still

ambiguous. Here too, many combinations of trends in birth prevalence and diagnostic factors can produce similar trends in cumulative incidence.

Ambiguity in estimation

How should one interpret a dataset that produces any one of the cumulative incidence curves illustrated in Fig 1? The figure represents synthetic data; some real-world data may be similar. Observed data might produce a curve resembling any one of the curves in the figure. An exponential curve with a coefficient of 0.1 fits all three plotted lines reasonably well. Does this represent a true increase in birth prevalence with a coefficient of 0.1? Does it result from an exponential increase in the effects of diagnostic factors, with no increase in birth prevalence? Perhaps a combination of both? The three similar cumulative incidence curves represent quite different possible explanations. How to distinguish which one is correct?

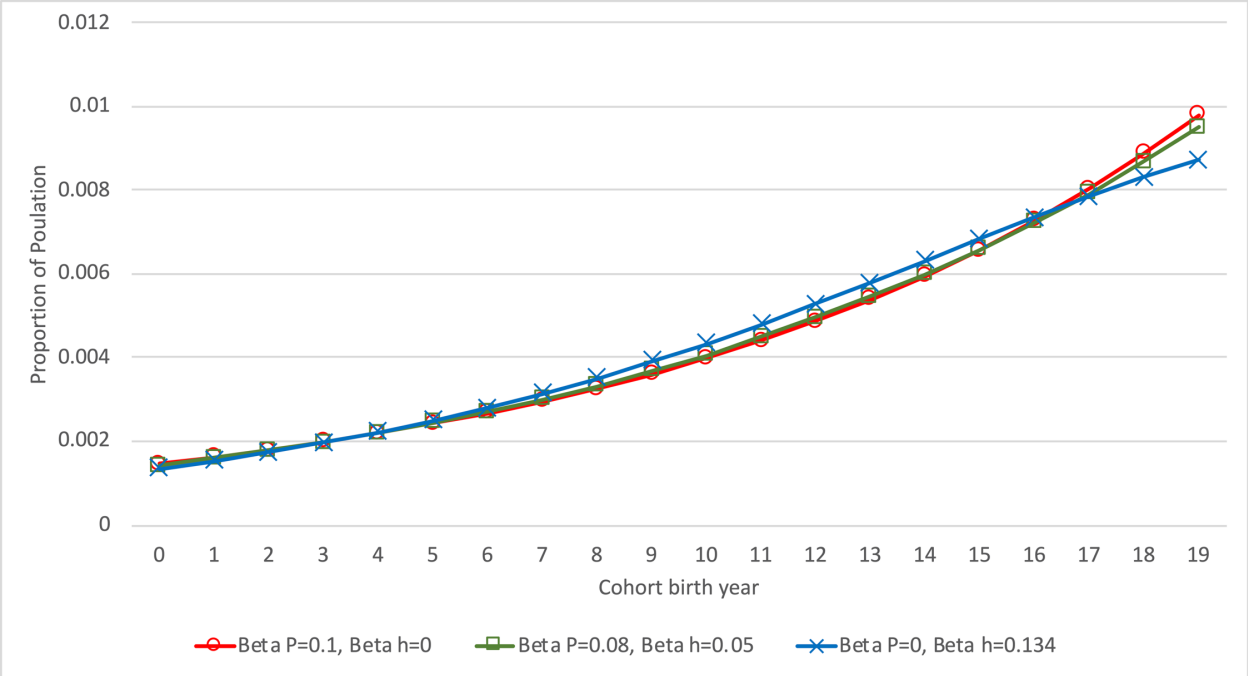

Fig 1. Example of cumulative incidence under three models. β_{P} is the coefficient for birth prevalence; β_{H} is the coefficient for the effect of diagnostic factors. Red line with circles represents $\beta_{\mathrm{P}} = 0.1, \beta_{\mathrm{H}} = 0$; green line with squares represents $\beta_{\mathrm{P}} = 0.08, \beta_{\mathrm{H}} = 0.5$; blue line with crosses represents $\beta_{\mathrm{P}} = 0, \beta_{\mathrm{H}} = 0.134$.

Fig 1 illustrates a hypothetical example of cumulative incidence to age ten over 20 consecutive cohorts. The legend lists the parameter value pairs for the three cases. β_P is the exponential coefficient of birth year prevalence, and β_h is the exponential coefficient of the effect of diagnostic factors by diagnostic year. The data generation process producing these data uses a survival process as detailed below. An Excel spreadsheet to generate all plots in this paper is available at OSF [28]. The variable h represents diagnostic pressure, the effect of diagnostic factors, which is equivalent to hazard, as explained in the section Significance of birth year and diagnostic year. While the three cumulative incidence curves appear similar, the age distributions of diagnoses are strikingly different, as Fig 2 shows. The remainder of this paper explains how modeling the age distribution of first diagnoses enables accurate and unambiguous estimation of the coefficients for birth prevalence and diagnostic factors.

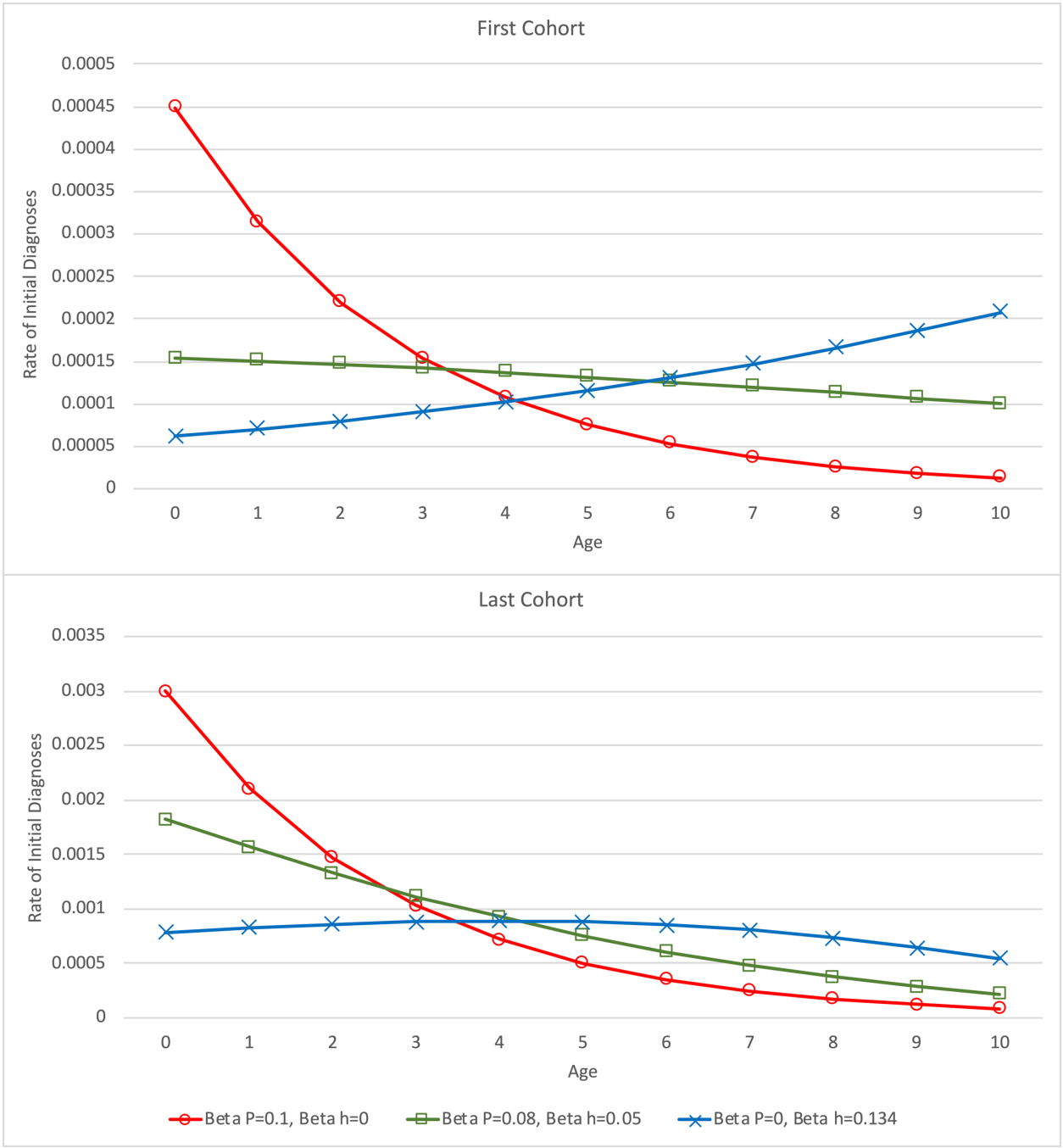

Fig 2. Distribution of diagnoses in the first and last cohorts under three models. β_{P} is the coefficient for birth prevalence; β_{H} is the coefficient for the effect of diagnostic factors. Red lines with circles represent $\beta_{\mathrm{P}} = 0.1, \beta_{\mathrm{H}} = 0$; green lines with squares represent $\beta_{\mathrm{P}} = 0.08, \beta_{\mathrm{H}} = 0.5$; blue lines with crosses represent $\beta_{\mathrm{P}} = 0, \beta_{\mathrm{H}} = 0.134$.

If the disorder's complete cause is in place at birth, then the incidence of the disorder is indistinguishable from birth prevalence. In contrast, diagnosis occurs later, if at all. The incidence of diagnosis represents the combination of birth prevalence and delays and omissions in diagnoses.

One might consider controlling for diagnostic factors over time, for example, via regression, but that is not sufficient to distinguish between alternative explanations. Diagnostic factors are a function of time, and birth year prevalence is also a function of time. As the Introduction states, controlling directly for diagnostic factors biases the estimates of the main effect, typically towards zero, citing Schisterman [19] and Rothman [2]. Schisterman recommends, "clearly stating a causal question to be addressed, depicting the possible data generating mechanisms using causal diagrams, and measuring indicated confounders." This paper directly addresses these issues.

Significance of birth year and diagnostic year

Diagnostic factors only affect the diagnosis of cases when those cases exhibit diagnosable symptoms, referred to as being eligible for diagnosis. Diagnostic pressure is the probability of diagnosing eligible undiagnosed cases, and it is an effect of the combination of all diagnostic factors. The Introduction lists examples.

Diagnostic pressure is equivalent to the hazard h in time-to-event or survival analysis. For each case of the disorder, the information resulting from diagnostic pressure consists of the time of initial diagnosis, that is, the diagnostic year. Diagnostic pressure has no observable effect before the diagnosis of each case, and none after the initial diagnosis since TTEPE considers only initial diagnoses. Hence, the effect of diagnostic pressure on the input data is a function of diagnostic year.

The directed acyclic graph (DAG) in Fig 3 illustrates the causal paths from birth year, diagnostic year, and age to diagnosis. Birth year drives etiologic (causal) factors, which produce the disorder and its symptoms. Diagnostic criteria determine whether each individual's symptoms qualifies them as a case, and criteria may change at specific diagnostic years. Symptoms may vary with age. Diagnostic year drives diagnostic factors, which form diagnostic pressure. Diagnosable symptoms and diagnostic pressure together produce each initial diagnosis.

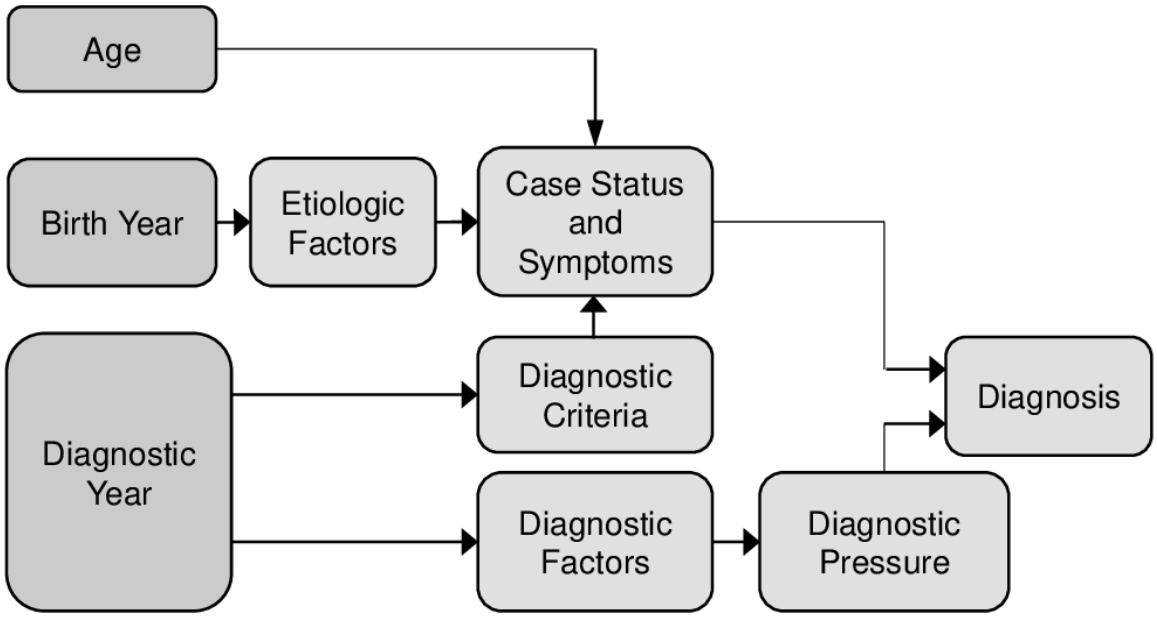

Fig 3. Directed Acyclic Graph Representing Year of Birth, Diagnostic Year and Age

Changes in diagnostic criteria can affect the threshold of symptoms that qualify case status. Criteria changes may change the proportion of the cohort classified as cases, i.e., the effective prevalence.

Development of the TTEPE method

The TTEPE method is based on the DAG of Fig 3 and modeling the age distribution of initial diagnoses. The method avoids the identification problem associated with age-period-cohort analysis, and it avoids the problem of inappropriate adjustment for diagnostic factors.

TTEPE is particularly applicable to disorders where case status is established by birth or by a known age, and diagnosable symptoms are present by some consistent age. It is also useful where cases develop diagnosable symptoms gradually over a range of ages.

Data sources suitable for TTEPE analysis provide rates of initial diagnoses by age or, equivalently, diagnostic year for each birth cohort. The rate is the count of initial diagnoses divided by the population of the cohort at the respective age.

TTEPE relies on these principles: for each birth cohort, the number of cases at risk of initial diagnosis decreases as cases are diagnosed, and diagnostic year is the time when diagnostic factors affect the probability of diagnosing cases exhibiting diagnosable symptoms.

First, consider the case where the diagnostic criteria do not change the effective prevalence over the interval of interest. The section Changes in criteria affecting prevalence examines the alternative.

The principle of finding the simplest model that fits the data, sometimes called Occam's Razor, could lead to the conclusion that birth year effects best explain the observed birth year prevalence trends. However, the most parsimonious model is not always the

best one [29]. Specifically, diagnostic pressure as a function of diagnostic year could potentially be an essential component for explaining observed diagnosis data.

TTEPE is an extension of established time-to-event methods. TTEPE simultaneously estimates the birth prevalence and diagnostic pressure functions by fitting a model to the rate of initial diagnoses at each data point. It models the age distribution of rates of initial diagnoses via a survival process as a function of birth year, diagnostic year (birth year plus age), and eligibility.

TTEPE introduces the concept of eligibility. A case is eligible for diagnosis if the individual has diagnosable symptoms and not eligible if the individual has not yet developed diagnosable symptoms. For each birth cohort, undiagnosed eligible individuals form the risk set of cases at risk of initial diagnosis. The size of the risk set is denoted R. At each age there is some probability of diagnosis of each case in the risk set. This probability is the diagnostic pressure h. At each age, newly diagnosed cases are removed from the risk set, and newly eligible cases are added to the risk set. For any given value of h, as R decreases or increases, the rate of initial diagnoses D changes accordingly. This process generates the modeled age distribution. The survival function S refers to cases that “survive” diagnosis at each age. If all cases were eligible from birth, S would equal R. More generally, however, some cases may initially be ineligible and become eligible as they age, so $R \leq S$. The prevalence P is the proportion

of the population that are or become cases. The eligibility factor E is the eligible proportion of P, $0 \leq E \leq 1$.

In contrast, in typical survival or time-to-event analysis, including Cox proportional hazards analysis [30], the initial size of the risk set is assumed to have a known value, for instance, an entire population or an entire sample. If the risk set R consisted of the entire population without subtracting diagnosed cases, the estimate $\widehat{h_t} = D_t/R_t$ would be equivalent to the population-based rate of diagnoses D at time t. If the disorder is rare, it makes little difference whether the risk set is the entire population or the undiagnosed portion.

Population-based rates of initial diagnoses D are observable while the other variables are not. Different values of diagnostic pressure h produce different values of D for a given value of prevalence P, as shown in the Illustrative example section.

Time-to-event analysis model

The analysis model enables estimation of the temporal trend of birth prevalence P over a range of cohorts, correctly adjusted for diagnostic pressure h. Both P and h can vary with time. P is a function of birth year, and h is a function of diagnostic year. Estimation of P adjusted for h requires estimating the time-based parameters of both P and h in the time-to-event model and specifying or estimating the eligibility function E.

Let $D_{BY,A}$ be the population-based rate of incident diagnoses, where BY is birth year, and A is age. The model generates predicted values $\widehat{D_{BY,A}}$. Modeling of proportions rather than counts accommodates changes in the population size of each cohort over time, e.g., due to in- and out-migration and deaths. Count values are useful for calculating p-values from chi-square goodness-of-fit measurements. Alternatively, the analysis could model counts directly.

Let h_{DY} be the diagnostic pressure at diagnostic year DY. $DY = BY + A$, subject to rounding, so h_{DY} is equivalent to $h_{BY,A}$. Let P_{BY} be the case prevalence of birth year cohort BY. Let $R_{BY,A}$ be the discrete risk set function of the population proportion of eligible cases at risk of initial diagnosis at age A for birth year BY. TTEPE uses R rather than a discrete survival function S to accommodate eligibility changing with age. Let E_A be the discrete eligibility function, the proportion of cases that are eligible at age A, bounded by $0 \leq E \leq 1$. At each age $A \geq 1$, $P \times (E_A - E_{A-1})$ is the incremental portion of prevalent cases added to R due to changes in eligibility. For simplicity, assume E_A increases monotonically, i.e., non-decreasing, meaning that cases do not lose eligibility before diagnosis.

Kalbfleisch [31] gives background on general time-to-event theory and equations.

Consider three scenarios, differing by the characteristics of E_A. Here we write h_{DY} as $h_{BY,A}$ to clarify the effect of A in $DY = BY + A$.

<u>Scenario: constant $E_A = 1$.</u> All cases are eligible from birth, so $E_A = 1$ for all values of A. This scenario is equivalent to standard time-to-event models that do not consider eligibility. For $A \geq 1$, $E_A - E_{A-1} = 0$. For the first year of age, $A = 0$, $R_{BY,0} = P_{BY}E_0 = P_{BY}$ and $D_{BY,0} = R_{BY,0}h_{BY,0} = P_{BY}h_{BY,0}$.

For $A = 1$, $R_{BY,1} = P_{BY} - D_{BY,0} = P_{BY} - P_{BY}h_{BY,0} = P_{BY}(1 - h_{BY,0})$ and $D_{BY,1} = R_{BY,1}h_{BY,1} = P_{BY}(1 - h_{BY,0})h_{BY,1}$.

For $A = 2$, $R_{BY,2} = R_{BY,1} - D_{BY,1} = P_{BY}\left(1 - h_{BY,0}\right) - P_{BY}\left(1 - h_{BY,0}\right)h_{BY,1} = P_{BY}\left(1 - h_{BY,0}\right)(1 - h_{BY,1})$ and $D_{BY,2} = R_{BY,2}h_{BY,2} = P_{BY}(1 - h_{BY,0})(1 - H_{BY,1})h_{BY,2}$.

Similarly, for $A = 3$,

$R_{BY,3} = P_{BY}(1 - h_{BY,0})(1 - h_{BY,1})(1 - h_{BY,2})$ and

$D_{BY,3} = P_{BY}(1 - h_{BY,0})(1 - h_{BY,1})(1 - h_{BY,2})h_{BY,3}$.

Generally, for $A \geq 1$,

$$R_{BY,A} = P_{BY}\prod_{a=0}^{A-1}(1 - h_{BY,a})$$

and

$$D_{BY,A} = P_{BY}\prod_{a=0}^{A-1}\left(1 - h_{BY,a}\right)h_{BY,A} \tag{1}$$

In all three scenarios in this paper, the survival function is:

$$S_{BY,A} = P_{BY} - \sum_{a=0}^{A-1} D_{BY,a} \tag{2}$$

The summation term is the cumulative incidence of initial diagnoses through age $A - 1$.

<u>Scenario: Increasing E_A.</u> $E_0 < 1$ and E_A increases monotonically with A. For $A = 0$, $R_{BY,0} = E_0 P_{BY}$ and $D_{BY,0} = E_0 P_{BY} h_{BY,0}$. For each $A \geq 1$, $R_{BY,A} = R_{BY,A-1} - D_{BY,A-1} + (E_A - E_{A-1})P_{BY}$. The incremental increase of E_A causes an incremental increase in $R_{BY,A}$. Then,

$$D_{BY,A} = R_{BY,A} h_{BY,A} = \left(R_{BY,A-1} - D_{BY,A-1}\right) h_{BY,A} + (E_A - E_{A-1}) P_{BY} h_{BY,A} \tag{3}$$

Equation (3) can be useful as a procedural definition. We can write equivalent expressions for $R_{BY,A}$ and $D_{BY,A}$ as sums of expressions similar to equation (1), where each summed expression describes the portion of P_{BY} that becomes eligible at each age according to E_A. For $A \geq 1$,

$$R_{BY,A} = \sum_{a=0}^{A-1} (E_a - E_{a-1}) P_{BY} \prod_{b=a}^{A-1} (1 - h_{BY,b})$$

$$D_{BY,A} = \sum_{a=0}^{A-1} (E_a - E_{a-1}) P_{BY} \prod_{b=a}^{A-1} \left(1 - h_{BY,b}\right) h_{BY,A} \tag{4}$$

where E_{-1} is defined to be 0. E_A can be defined parametrically or non-parametrically.

Scenario: Plateau E_A. E_A increases from $E_0 < 1$ and plateaus at $E_A = 1$ for $A \geq AE$, where AE is the age of complete eligibility, $AE < M$, and M is the maximum age included in the analysis. Equation (3) applies, noting that for $A > AE, (E_A - E_{A-1}) = 0$. Equivalently, combine equation (2) with the fact that $E_{AE} = 1$ to obtain $R_{AE} = S_{AE} = P_{BY} - \sum_{a=0}^{AE-1} D_{BY,a}$, so

$$D_{BY,AE} = R_{BY,AE} h_{BY,AE} = S_{BY,AE} h_{BY,AE} = (P_{BY} - \sum_{a=0}^{AE-1} D_{BY,a})\, h_{BY,AE} \tag{5}$$

and for $A > AE$,

$$R_{BY,A} = S_{BY,A} = (P_{BY} - \sum_{a=0}^{AE-1} D_{BY,a}) \prod_{b=AE}^{A-1} (1 - h_{BY,b})$$

$$D_{BY,A} = (P_{BY} - \sum_{a=0}^{AE-1} D_{BY,a}) \prod_{b=AE}^{A-1} \left(1 - h_{BY,b}\right) h_{BY,A} \tag{6}$$

The scenario of increasing E_A is a general formulation and may not be needed in practice. The plateau E_A scenario may be appropriate when external information, such as the definition of the disorder, indicates the value of AE, or when investigators specify AE based on estimates of E_A found using equation (4). Equations (5) and (6) do not model E_A nor $D_{BY,A}$ for $A < AE$. Rather, they use the empirical values of $D_{BY,A}$ for $A < AE$.

Prevalence, cumulative incidence and censoring

The case prevalence in each cohort is the cumulative incidence of initial diagnoses through the last age of follow-up plus the censored portion. This assumes that any difference in competing risks between cases and non-cases in the age range analyzed is small enough to be ignored. This assumption is consistent with Hansen [15]. If the rate of deaths of cases before initial diagnosis exceeds that of the entire population of the cohort at the same ages, that excess would constitute a competing risk and would reduce the estimated prevalence accordingly.

In all three scenarios of E_A, we can express P as a function of S and the cumulative incidence $CI = \sum_{a=0}^{A-1} D_{BY,a}$ for $A > 0$, by rearranging equation (2) as $P = S_A + CI_{A-1}$. Assuming that eligibility at the last age of follow-up $E_M = 1$, $S_M = R_M$. Then, $P = R_M + CI_{M-1}$ and $D_M = R_M h_M$. The censored proportion is $S_{M+1} = S_M - D_M$, which is equivalent to $S_{M+1} = R_M - R_M h_M = R_M(1 - h_M)$. After estimating the model parameters, the estimated censored proportion is $\widehat{S_{M+1}} = \widehat{R_M}(1 - \widehat{h_M})$.

Illustrative example

Fig 4 illustrates an example according to the plateau E_A scenario showing the relationships between prevalence, diagnosis rates, the survival function, and cumulative incidence CI with two different values of diagnostic pressure h = 0.1 and h = 0.25 and prevalence P = 0.01. In this example, $E = 1$ for $A \geq AE = 3$ and h takes on one of two constant values. The value of h determines the shapes of these functions vs. age. This

example shows constant values of h purely for clarity, not as an assumption nor a limitation of TTEPE.

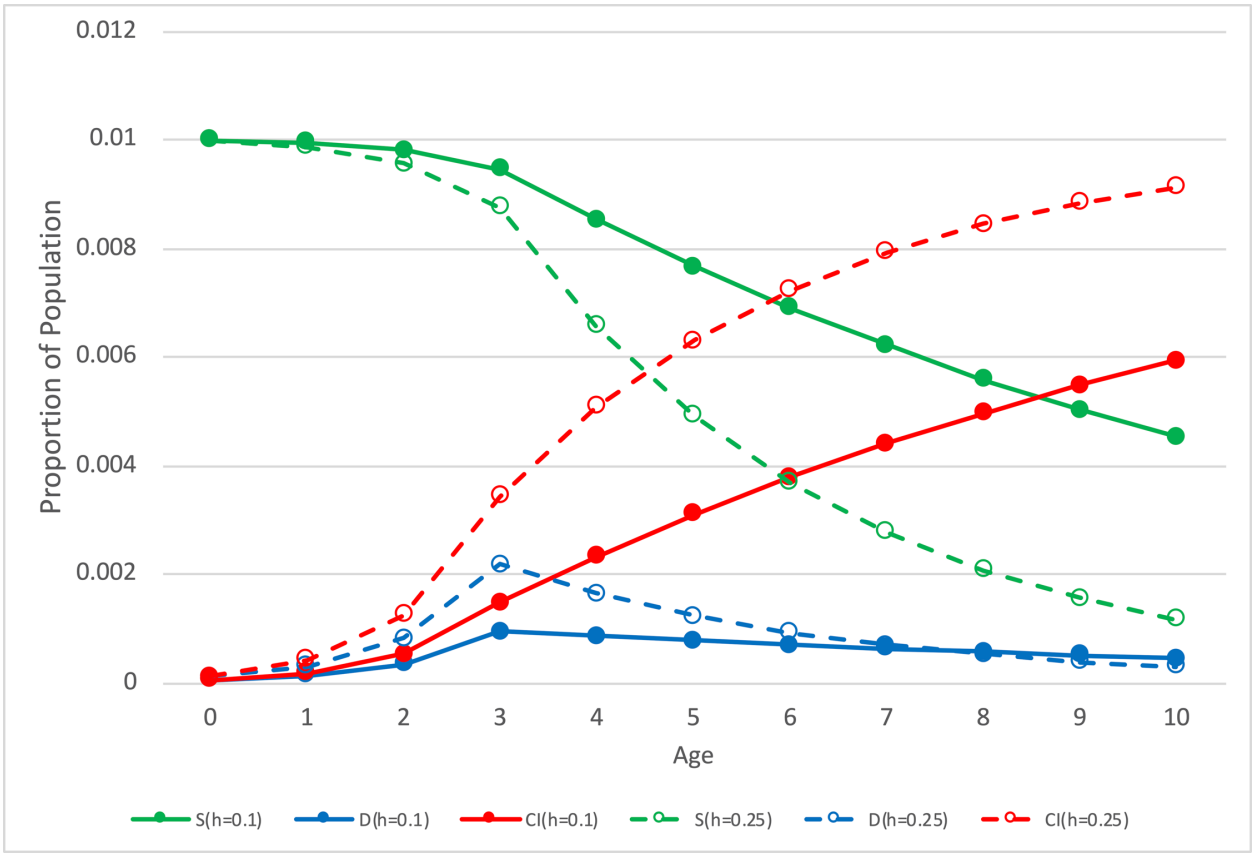

Fig 4. Example of a survival process for two values of diagnostic pressure h. The green lines S denote survival, the blue lines D denote the rate of diagnoses, and the red lines CI denote cumulative incidence. The solid lines represent h=0.1, and the dotted lines represent h=0.25.

As cases are diagnosed, S decreases and CI increases. R is not shown; $R = S$ for $A \geq AE = 3$. Only D is observable.

Assumptions

Several baseline assumptions enable TTEPE analysis. Some assumptions may be relaxed, as discussed below.

1. The eligibility function E_A under consistent diagnostic criteria is consistent across cohorts.
2. The diagnostic pressure applies equally to all eligible undiagnosed cases at any given diagnostic year.
3. The case prevalence under consistent diagnostic criteria within each cohort is constant over the range of ages included in the analysis.
4. Case status is binary according to the applicable diagnostic criteria.
5. The discrete-time interval (e.g., one year) is small enough that the error introduced by treating the variable values as constant within each interval is negligible.
6. No false positives.
7. Data represent truly initial diagnoses.
8. Any difference in competing risks between cases and non-cases in the age range analyzed is small enough to be ignored.

The assumption of a consistent eligibility function means that cases develop diagnosable symptoms as a function of age, and that function is the same for all cohorts

under consistent diagnostic criteria. The section Changes in criteria affecting prevalence discusses a separate effect that might make the eligibility function appear inconsistent.

Estimating parameters

TTEPE performs a non-linear regression that estimates the parameters of a model of $D_{BY,A}$ using general-purpose optimization software. The model is based on equations (1) through (6) selected based on the eligibility scenario. The model produces estimates $\widehat{D_{BY,A}}$ from the parameters and independent variables, and the software finds the parameter values that minimize a cost function cost($D, \widehat{D}$). One suitable implementation of optimization software in the Python language is the curve_fit() function in the SciPy package (scipy.optimize.curve_fit in SciPy v1.5.2). Its cost function is $(D - \widehat{D})^2$, so it minimizes the sum of squared errors. Python software to perform this regression and the simulations described below is available at OSF [28].

Investigators should choose which model equation to use based on knowledge or estimates of the eligibility function E_A. The constant E_A scenario and equation (1) assume that all cases are eligible from birth, which may not be valid for some disorders. The validity of the assumption that all cases are eligible by a known age AE, i.e., the plateau E_A scenario and equation (6), may be supported by either external evidence, e.g., the definition of the disorder, or estimation of E_A. The least restrictive approach of the increasing E_A scenario uses equation (4) to estimate E_A. Non-parametric estimates $\widehat{E_A}$ can inform a choice of a parametric form of E_A. The value of E_A at the maximum age studied M should be set to 1 to ensure the estimates are identifiable. If $E_A = 1$ for all $A \geq$

AE, that fact and the value of AE should be apparent from estimates $\widehat{E_A}$, and the plateau E_A scenario applies.

Investigators should choose forms of P_{BY} and $h_{BY,A}$ appropriate to the dataset. Linear, first-order exponential, second-order exponential or non-parametric models may be appropriate. Graphical and numeric model fit combined with degrees of freedom can guide the optimum choice of a well-fitting parsimonious model.

TTEPE preferably estimates P and h simultaneously over a series of cohorts, utilizing data points from all cohorts, thereby enabling well-powered estimation and flexible model specification. Alternatively, under some conditions, it may be possible to estimate h in a single cohort and estimate P based on $\hat{h}$.

Suppose the population proportion of cases represented in the data is unknown for all cohorts. In that case, estimates of the absolute prevalence, or the intercept, may be underestimated by an unknown scale factor. If that proportion is known for at least one cohort, we can use it to calibrate the intercept. Proportional changes in prevalence between cohorts are unaffected by underestimation of the intercept. If the population proportion of cases included in the sample changes over time, that change reflects changing diagnostic factors, and the estimated parameters of h automatically represent such changes.

Changes in criteria affecting prevalence

Changes in diagnostic criteria could potentially affect rates of initial diagnoses by changing the effective prevalence. This mechanism is distinct from diagnostic pressure. Changes in criteria may change the effective prevalence within a cohort, without affecting symptoms or etiology, by including or excluding as cases some portion of the cohort population compared to prior criteria. A criteria change may change the effective prevalence of the entirety of any cohort where the birth year is greater than or equal to the year the change took effect. For birth years before the year of criteria change, a change in criteria that changes the effective prevalence causes an increase or decrease in the size of the risk set R starting at the diagnostic year the change took effect. Generally, diagnostic criteria should be given in published documents, such that changes in criteria correspond to effective dates of new or revised specifications.

Let $\{CF_{cy}\}$ be the set of criteria factors that induce a multiplicative effect on effective prevalence due to criteria changes that occurred at criteria years $\{cy\}$ after the first DY included in the study. P_{BY} is the prevalence of cohort BY before the effect of any of $\{CF_{cy}\}$. For each cohort BY, the effective prevalence $EP_{BY,A}$ at age A is

$$EP_{BY,A} = P_{BY} \prod_{cy \leq (BY+A)} CF_{cy} \tag{7}$$

where CF_0, the value in effect before the first DY in the study, equals 1. The combination of P_{BY} and the effects of all $\{CF_{cy} | cy \leq BY + A\}$ determines the final effective prevalence of each cohort.

For a given BY and increasing A, $BY + A$ crossing any cy causes a step-change in the effective prevalence EP. Using a general formulation of eligibility E_A, per the increasing E_A scenario and equation (3), and for clarity substituting $BY + A$ for DY, we obtain the following. For $A = 0$, $R_{BY,0} = E_0 EP_{BY,0}$ and $D_{BY,0} = E_0 EP_{BY,0} h_{BY,0}$. For $A \geq 1$,

$$R_{BY,A} = R_{BY,A-1} - D_{BY,A-1} + E_A(EP_{BY,A} - EP_{BY,A-1}) + (E_A - E_{A-1})EP_{BY,A}$$

and

$$D_{BY,A} = [R_{BY,A-1} - D_{BY,A-1} + E_A(EP_{BY,A} - EP_{BY,A-1}) + (E_A - E_{A-1})EP_{BY,A}]h_{DY} \qquad (8)$$

The term $EP_{BY,A} - EP_{BY,A-1}$ represents the change in the effective prevalence EP when $BY + A$ crosses one of $\{cy\}$. As each CF_{cy} takes effect at $cy = DY = BY + A$, the newly effective CF_{cy} changes $EP_{BY,A}$ and R in all BY cohorts where cy corresponds to an age A in the range of ages studied. These changes in R affect the rates of initial diagnoses D. For cohorts born after cy, CF_{cy} applies to all ages.

The parameters of P_{BY} quantify the birth year prevalence controlled for diagnostic criteria changes, which are represented by $\{CF_{cy}\}$. In other words, P_{BY} is the cohort prevalence that would have occurred if the initial criteria had been applied at all diagnostic years included in the study.

To estimate the parameters, use a software model of equation (8) with optimization software, as described in the previous section.

Potential violations of assumptions

Suppose a dataset represents a non-homogeneous set of cases with different effective values of diagnostic pressure h applying to different unidentified subgroups at the same DY. That would violate the assumption that the diagnostic pressure applies equally to all eligible undiagnosed cases at any given DY. Cases may have differing degrees of symptom severity, and more severe symptoms may result in earlier diagnosis [32], implying greater diagnostic pressure. Fig 5 illustrates this situation. The figure illustrates constant values of h purely for clarity, not as an assumption nor a limitation. If data represent unidentified subgroups with differential diagnostic pressure, the distribution of diagnoses is a sum of distributions with different values of h. Such a sum of distributions with different values of h may impair fit with a model that assumes homogeneous h. For data where $E = 1$ for $A \geq AE$ (plateau E_A), adjusting the assumed value of AE used in estimation, called AE^*, may mitigate such errors, as shown in the Simulation study section. Stratified estimation using subgroup data, if available, can avoid the issue of unidentified non-homogenous subgroups.

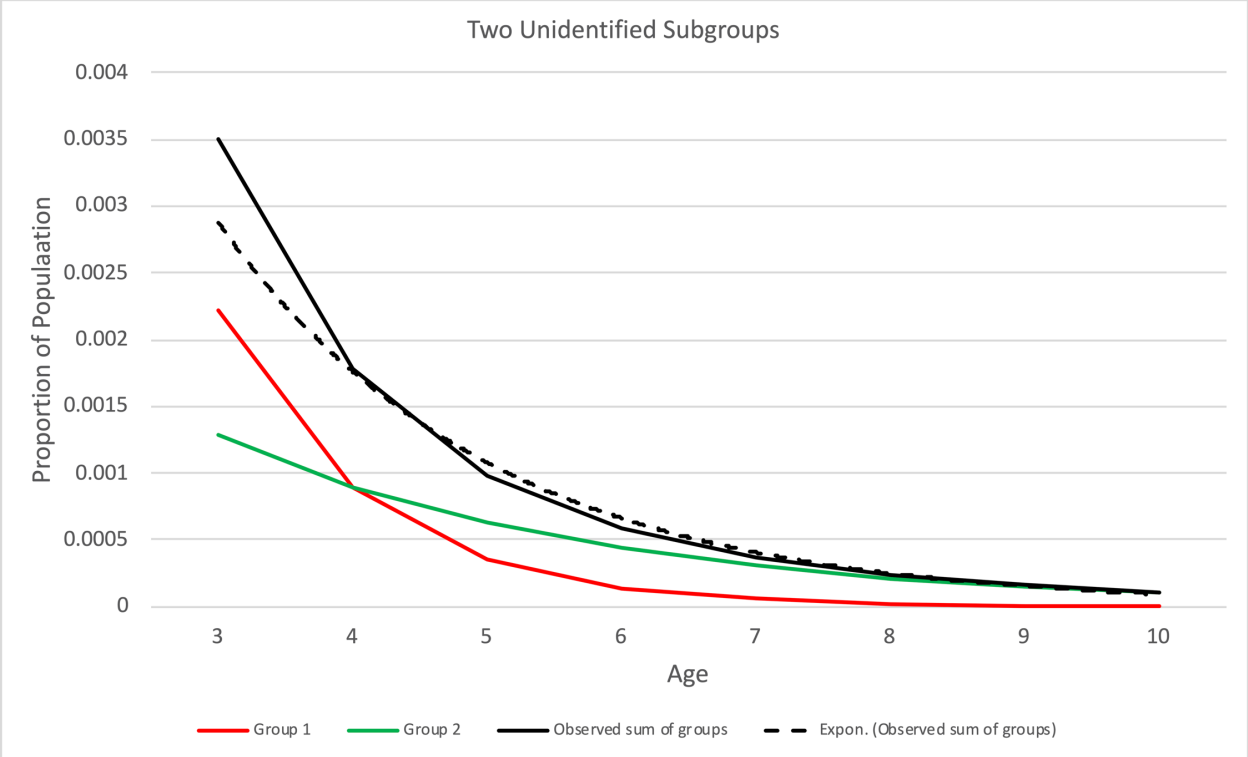

Fig 5. Example where observed diagnosis rates represent two unidentified subgroups with different values of diagnostic pressure. The red and green lines represent rates of diagnosis of the two subgroups. The solid black line shows the aggregate diagnosis rates. The dotted line shows the exponential fit to the aggregate diagnosis rates. The age of eligibility $AE = 3$ in this example.

Any imbalance of case prevalence between in-migration and out-migration to and from the region defining the population over the study period would violate the assumption of constant prevalence within each cohort.

If some in-migrating cases were diagnosed before in-migration and their subsequent re-diagnoses in the study region were labeled as initial diagnoses, that would violate the

assumption of truly first diagnoses. Such an effect would be most evident at greater ages after the diagnosis of most cases. Bounding the maximum age studied M to a modest value, sufficient to capture most initial diagnoses, can minimize any resulting bias.

Apart from subgroups with non-homogeneous h, it is theoretically possible for h to have different effective values for cases of different ages with the same symptom severity at the same DY. Such an effect would represent an age bias in diagnostic pressure, independent of symptom severity. If there is a reason to suspect such an age bias, investigators can add an age parameter to h in the model and estimate its parameters.

For datasets where diagnosis follows best practices using gold-standard criteria, the lack of false positives may be a fair assumption. It would be difficult to discover any false positives in that case. Where diagnosis uses a less precise process, some false positives might occur. For example, diagnosticians might tend to produce a positive diagnosis of individuals who do not meet formal diagnostic criteria, perhaps under pressure from the patient or parents, or to facilitate services for the individual. In scenarios where the rate of false positives is significant, the age distribution relative to the rates of true diagnoses may be important. If false positives are uniformly distributed over the age range studied, they would cause a constant additive offset to the rates of diagnoses. True case diagnoses should be more common at younger ages, and less common as the risk pool is depleted, so false positives may be relatively more evident

at older ages. If false positives are more common at older ages, their effect may be even more obvious.

Model fit

To ensure robust conclusions, investigators should test the model fit to ascertain both model correctness and parameter estimation accuracy. The model fits well if summary measures of the error are small and individual point errors are unsystematic and small [33]. One can examine the fit both graphically and numerically. Plots of $D_{BY,A}$ vs. $\widehat{D_{BY,A}}$ at all ages for individual cohorts can illuminate any issues with fit, which might occur at only some cohorts or ages. Visualization of the model vs. data can expose aspects of the data that might not fit well in a model with few parameters, possibly suggesting a higher-order model or semi-parametric specifications.

If the model uses an assumed age of complete eligibility AE^* that differs from the true value of AE represented by the data, model fit may be impaired, particularly if $AE^*<AE$. As the Simulation study section shows, setting $AE^*<AE$ can result in estimation errors. Setting $AE^*>AE$ tends not to impair model fit and may improve it in the case of non-homogeneous subgroups; see Fig 5. The presence of non-homogeneous subgroups may be evident from examining model fit.

The chi-square test statistic, applied to the overall model, individual cohorts, and single ages across cohorts, is a numerical approach to assess absolute model fit. The p-value

associated with the chi-square statistic utilizes observed and expected count values, not proportions. The p-value incorporates the effect of the number of parameters in the model via the degrees of freedom.

Simulation study

We tested the TTEPE method via a simulation study where we know the ground truth of all parameters, following the recommendations in Morris [34]. There are six pairs of values of β_h and β_P, each ranging from 0 to 0.1 in steps of 0.02, and each pair sums to 0.1. In one parameter set, the prevalence increases as $e^{0.1\times BY}$ and diagnostic pressure is constant; in another parameter set, the prevalence is constant and diagnostic pressure increases as $e^{0.1\times DY}$; and the other four parameter sets represent various rates of change of both variables. In all cases, $P = 0.01$ at the final BY, $h = 0.25$ at the final DY, $AE = 3$, $M = 10$, and there are 20 successive cohorts. These simulations assume the investigators know the correct value $AE^* = 3$ from either knowledge of the disorder or estimation of E_A. The study synthesized each data model as real-valued proportions without sampling and with binomial random sample generation of incident diagnoses. For sampling, the population of each cohort is a constant 500,000. Monte Carlo simulation of parameter estimation bias and model standard error (SE) used 1000 iterations of random data generation for each set of parameters. Parameter estimation is as described above, implemented using the Python SciPy curve_fit() function.

Table 1 shows results using real-valued proportions without sampling, isolating the estimation process from random sampling variations. It shows the bias in estimating each of the four model parameters for each of the six combinations of β_h and β_P. The biases are minimal; the greatest bias magnitude in $\widehat{\beta_P}$ occurs with β_P = 0 and β_h = 0.1 and is on the order of 10^{-10}.

Table 1. Simulation results of parameter optimization using real-valued proportions with no sampling.

True Parameters		Bias $\widehat{P}$ at final BY	Bias $\widehat{\beta_P}$	Bias $\widehat{h}$ at final DY	Bias $\widehat{\beta_h}$
β_P	β_h				
0.1	0	5.9E-12	8.9E-11	-5.5E-10	-1.8E-10
0.08	0.02	0	0	-2.8E-17	-1.4E-17
0.06	0.04	-1.7E-18	0	1.1E-16	1.4E-17
0.04	0.06	3.5E-18	1.4E-17	-5.6E-17	-6.9E-18
0.02	0.08	0	3.5E-18	5.6E-17	-1.4E-17
0	0.1	-4.7E-11	-6.6E-10	2.2E-9	7.7E-10

P = 0.01 at the final BY, h =0.25 at the final DY, AE^*=AE=3, M = 10, and there are 20 successive cohorts.

β_P, β_h are coefficients for prevalence and hazard, respectively.

P, prevalence; BY, birth year; DY, diagnostic year.

Table 2 gives results from Monte Carlo analysis of the same parameter sets where the data use binomial sampling. It shows the bias and model SE of each parameter for each parameter set. The bias of the primary parameter $\widehat{\beta_P}$ remains small, on the order of 10^{-5} or 10^{-6}. The SE is relevant when considering sampling, and it shows the effect of sampling compared to Table 1.

Table 2. Simulation results of parameter optimization using Monte Carlo with binomial sampling, 1000 iterations.

True Parameters		$\widehat{P}$ at final BY		$\widehat{\beta_P}$		$\widehat{h}$ at final DY		$\widehat{\beta_h}$	
β_P	β_h	Bias	SE	Bias	SE	Bias	SE	Bias	SE
0.1	0	-2.0E-6	1.0E-4	-2.0E-5	0.0013	3.3E-5	0.0070	-5.4E-6	0.00191
0.08	0.02	-2.6E-6	1.1E-4	-3.2E-5	0.0012	1.5E-4	0.0072	4.4E-5	0.00191
0.06	0.04	7.8E-6	1.2E-4	2.7E-5	0.0013	-4.4E-4	0.0079	-1.2E-4	0.00207
0.04	0.06	6.5E-5	0.0015	6.5E-5	0.0015	-5.8E-4	0.0085	-1.6E-4	0.00224
0.02	0.08	-2.0E-6	1.6E-4	-9.8E-6	0.0016	4.4E-4	0.0086	9.4E-5	0.00227
0	0.1	4.5E-6	1.8E-4	7.1E-6	0.0017	2.0E-4	0.0094	2.7E-5	0.00234

Population of each cohort = 500,000. P=0.01 at the final BY, h=0.25 at the final DY, AE^*=AE=3, M = 10, and there are 20 successive cohorts.

β_P, β_h are coefficients for prevalence and hazard, respectively.

P, prevalence; BY, birth year; DY, diagnostic year.

Table 3 gives results where estimation uses an assumed value AE^* that, in some cases, does not match the true value of $AE = 3$ represented by the data. Synthesis uses one homogenous group with consistent h at each value of DY. Estimation using $AE^* = 2$ results in substantial estimation errors and model misfit that is obvious from plots of data vs. model (not shown). Estimation using $AE^* = 3$, $AE^* = 4$, or $AE^* = 5$ produces accurate results, with slightly more error where $AE^* = 5$. Plots show that the model fits well in all three cases (not shown). The choice of AE^* is not critical as long as $AE^* \geq AE$. These data use real-valued proportions to avoid confusing model mismatch with sampling effects.

Table 3. Comparison of the effect of the choice of assumed AE^* vs. true value of $AE = 3$, with one homogeneous group of cases.

AE^* used in estimation	Bias $\widehat{P}$ at final BY	Bias $\widehat{\beta_P}$	Bias $\widehat{h}$ at final DY	Bias $\widehat{\beta_h}$
2	0.002	-0.019	-0.0096	0.036
3	5.9E-12	8.9E-11	-5.5E-10	-1.8E-10
4	-4.4E-12	-6.6E-11	4.8E-10	1.64E-10
5	1.5E-11	2.2E-10	-1.85E-9	-7.18E-10

AE, age of complete eligibility. True values: $\beta_P = 0.1$, $\beta_h = 0$, $P = 0.01$ at the final BY, $h = 0.25$ at the final DY, $AE = 3$. Maximum age $M = 10$. 20 successive cohorts. Diagnostic pressure is consistent across cases at each DY. Simulation uses real values, no sampling.

Table 4 shows results with an intentional mismatch between estimation assuming one homogeneous group and data representing two subgroups with different values of h, illustrated in Fig 5. Note the visible error of the exponential fit to the data at age = 3 and a good fit for age > 3. In this synthetic dataset, the two subgroups are of equal size, and the true value of h in one group is twice that of the other. This information is not known to the estimation, and the data do not indicate subgroup size nor membership. In the worst case, estimation uses $AE^* = AE = 3$, and the $\widehat{\beta_P}$ bias is 0.001, which is 1% of the actual value of 0.1. This error is due to the subgroups having different hazards, which

are not accounted for in the estimation. When using AE^* = 4 or AE^* = 5, the $\widehat{\beta_P}$ bias becomes 6×10^{-4} or less, and model fit is improved (not shown).

Table 4. Comparison of the effect of the choice of assumed AE^* vs. true value of AE = 3, with two unidentified subgroups with different hazards, mismatched to analysis.

AE^* used in estimation	Bias $\widehat{P}$ at final BY	Bias $\widehat{\beta_P}$	Bias $\widehat{h}$ at final DY	Bias $\widehat{\beta_h}$
3	-4.3E-4	0.001	0.0018	-0.002
4	-3.8E-4	6.1E-4	-0.004	-0.0016
5	-3.3E-4	3.5E-4	-0.0097	-0.0011

AE, age of complete eligibility. True values: β_P = 0.1, β_h = 0, P = 0.01 at the final BY, h =0.25 at the final DY, AE = 3. Two equal-sized groups of cases where the diagnostic pressure h of one group is twice that of the other, while the estimation assumes one homogeneous group. Maximum age M = 10. 20 successive cohorts. Simulation uses real values, no sampling.

Discussion

Readers may suspect that the estimates are unidentified, i.e., not unique, due to possible interaction between age, diagnostic year, and birth year, such that estimates may be biased even if the model fit is excellent. While that concern is appropriate for analytical methods that assume an age distribution, ignore it, or estimate it

inappropriately, including age-period-cohort methods, TTEPE avoids that problem and produces uniquely identified estimates. TTEPE models the age distribution of initial diagnoses as a non-linear function of birth year, diagnostic year, and other variables.

Keep in mind that the eligibility function E_A is different from the age distribution of diagnoses; E_A is an attribute of the disorder under study.

This paper states the assumptions that underlie TTEPE analysis. The DAG of Fig 3 illustrates the assumed causal paths from birth year, diagnostic year, and age, including the set of time-varying diagnostic factors and the effect of changes in criteria on effective prevalence. The DAG and associated analysis appear to cover all plausible mechanisms to explain observed trends in rates of initial diagnoses.

TTEPE provides accurate estimates of prevalence parameters with a strong power to detect small differences. The Monte Carlo simulation study in Table 2 shows a magnitude of bias of the prevalence coefficient $\widehat{\beta_P}$ not exceeding 6.5×10^{-5} or 0.0065% per year. The model SE of $\widehat{\beta_P}$ ranges from 0.0012 to 0.0017, where the true β_{P} ranges from 0 to 0.1. Using the largest observed SE and $1.96 \times SE$ as a threshold for 95% confidence intervals, the method can detect differences in β_{P} of 0.0033, i.e., 0.33% per year. Investigators can expect similar performance for real-world datasets that meet the baseline assumptions and have characteristics comparable to the simulated data. The population size and prevalence affect the SE. Note that in the simulation study, there

are 20 cohorts and 11 ages (0 through 10), so there are 220 data points. Each data point is an independent binomial random sample. The analysis estimates four parameters that define the curves that fit the data. The large number of independent data points and the small number of model parameters help to produce the small bias and model SE. If the population of each cohort or the prevalence was substantially smaller, or if the number of parameters was greater, we would expect the bias and SE to be larger. These could occur with small geographic regions, very rare disorders, or higher-order or semi-parametric models, respectively.

TTEPE is useful for answering some important questions, such as the actual trend in case prevalence over multiple birth cohorts of disorders such as autism and intellectual disability, as described in Elsabbagh [17] and McKenzie [26] respectively. Accurate trend estimates can inform investigation into etiology. Where datasets include appropriate covariates, stratified analysis can estimate the relationships between various population characteristics and trends in true case prevalence and diagnostic factors. Example covariates include sex, race, ethnicity, socio-economic status, geographic region, parental education, environmental exposure, genetic profile, and other potential factors of interest.

It may be feasible to extend TTEPE to disorders where the time scale starts at some event other than birth. For example, the time origin might be the time of completion of a sufficient cause, and various outcomes may serve as events of interest. It is important

to ensure that the eligibility function with respect to the time origin is consistent across cohorts.

Investigators may utilize domain knowledge to inform specialized analyses. For example, they may incorporate knowledge of mortality rates and standardized mortality ratios, rates of recovery from the condition before diagnosis, or the characteristics of migration in and out of the study region.

Acknowledgments

The author thanks Dr. Lu Tian for his expert advice on survival analysis methods; Dr. Lorene Nelson and Dr. Kristin Sainani for guidance on my thesis which was the genesis of this project and for comments on this paper; Dr. Michael Sigman and Dr. Larry Tang for their thoughtful reviews of the paper; and the Stanford Biomedical Data Science team for their project reviews and insightful comments.

References

1. Szklo M, Nieto FJ. Epidemiology Beyond the Basics. 1st ed. Burlington (MA): Jones & Bartlett; 2014.
2. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. 3rd ed. Philadelphia: Wolters Kluwer; 2008.

3. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018;392:1789–858. Suppl 1.

4. Baxter AJ, Brugha TS, Erskine HE, Scheurer RW, Vos T, Scott JG. The epidemiology and global burden of autism spectrum disorders. Psychol Med. 2015;45(3):601–613.

5. Centers for Disease Control. Prevalence of Autism Spectrum Disorders — Autism and Developmental Disabilities Monitoring Network, Six Sites, United States, 2000. MMWR Surveillance Summaries. 2007; 56(SS01);1-11. Available from: https://www.cdc.gov/mmwr/index.html.

6. Centers for Disease Control. Prevalence of Autism Spectrum Disorders — Autism and Developmental Disabilities Monitoring Network, Six Sites, United States, 2002. MMWR Surveillance Summaries. 2007; 56(SS01);12-28.

7. Centers for Disease Control. Brief Update: Prevalence of Autism Spectrum Disorders — Autism and Developmental Disabilities Monitoring Network, United States, 2004. MMWR Surveillance Summaries. 2009; 58(SS-10);21-24.

8. Centers for Disease Control. Prevalence of Autism Spectrum Disorders — Autism and Developmental Disabilities Monitoring Network, United States, 2006. MMWR Surveillance Summaries. 2009; 58(SS-10);1-20.

9. Centers for Disease Control. Prevalence of Autism Spectrum Disorders — Autism and Developmental Disabilities Monitoring Network, 14 Sites, United States, 2008. MMWR Surveillance Summaries. 2012;61(SS-3):1-19.

10. Centers for Disease Control. Prevalence of Autism Spectrum Disorders — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010. MMWR Surveillance Summaries. 2014;63(SS-2):1-21.

11. Centers for Disease Control. Prevalence of Autism Spectrum Disorders — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. MMWR Surveillance Summaries. 2018;65(13):1-23.

12. Centers for Disease Control. Prevalence of Autism Spectrum Disorders — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. MMWR Surveillance Summaries. 2018;67(6):1-23.

13. Centers for Disease Control. Prevalence of Autism Spectrum Disorders — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2016. MMWR Surveillance Summaries. 2020;69(4):1-12.

14. Croen LA, Grether JK, Hoogstrate J, Selvin S. The Changing Prevalence of Autism in California. J Autism Dev Disord. 2002;32(3):207-215.

15. Hansen SN, Overgaard M, Andersen PK, Parner ET. Estimating a population cumulative incidence under calendar time trends. BMC Med Res Methodol. 2017;17:7.

16. Nevison C, Blaxill M, Zahorodny W. California Autism Prevalence Trends from 1931 to 2014 and Comparison to National ASD data from IDEA and ADDM. J Autism Dev Disord. 2018; (doi.org/10.1007/s10803-018-3670-2). Suppl S1.

17. Elsabbagh M, Divan G, Koh Y-J, Kim YS, Kauchali S, Marcin C, et al. Global prevalence of autism and other pervasive developmental disorders. Autism Research. 2012;5:160–179.

18. Campbell CA, Davarya S, Elsabbagh M, Madden L, Fombonne E. Prevalence and the Controversy. In: Matson JL, Sturmey P, editors. International Handbook of Autism and Pervasive Developmental Disorders. New York: Springer; 2011 pp. 25-35.

19. Schisterman EF, Cole SR, Platt RW. Overadjustment Bias and Unnecessary Adjustment in Epidemiologic Studies. Epidemiology. 2009;20(4):488-495.

20. Keyes KM, Susser E, Cheslack-Postava K, Fountain C, Liu K, Bearman PS. Cohort effects explain the increase in autism diagnosis among children born from 1992 to 2003 in California. Int J Epidemiol. 2012;41(2):495-503

21. Spiers N. Cohort effects explain the increase in autism diagnosis among children born from 1992 to 2003 in California [letter]. Int J Epidemiol. 2013;42:1520–1521.

22. King M, Bearman P. Diagnostic change and the increased prevalence of autism. Int J Epidemiol. 2009; 38:1224–1234.

23. Rodgers WL. Estimable Functions of Age, Period, and Cohort Effects. Am Sociol Rev. 1982;47(6):774-787.

24. O'Brien RM. Age-Period-Cohort Models. Boca Raton (FL): CRC Press; 2015.

25. MacInnis AG. Autism Prevalence Trends by Birth Year and Diagnostic Year: Indicators of Etiologic and Non-Etiologic Factors – an Age Period Cohort Problem [thesis]. Stanford (CA): Stanford University; 2017 DOI: 10.13140/RG.2.2.11821.59360. Available from: https://www.researchgate.net/publication/322724736_Thesis_Autism_Prevalence_Trends_by_Birth_Year_and_Diagnostic_Year_Indicators_of_Etiologic_and_Non-Etiologic_Factors_-_an_Age_Period_Cohort_Problem.

26. McKenzie K, Milton M, Smith G, Ouellete-Kuntz H. Systematic Review of the Prevalence and Incidence of Intellectual Disabilities: Current Trends and Issues. Cur Dev Disord Rep. 2016;3:104-115.

27. Fombonne E. Epidemiology of pervasive developmental disorders. Pediatr Res. 2009;65(6):591-598.

28. MacInnis AG. Time-to-event Prevalence Estimation TTEPE [software]. 2020. OSF repository. Available from: https://doi.org/10.17605/OSF.IO/WPNKU.

29. Findley DF. Counterexamples to Parsimony and BIC. Ann Inst Stat Math. 1991;43(3):505-514.

30. Cox DR. Regression Models and Life Tables. J R Stat Soc Series B Stat Methodol. 1972;34(2):187-220.

31. Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data. 2nd ed. Hoboken (NJ): Wiley; 2002.

32. Hyman SL, Levy SE, Myers SM, AAP COUNCIL ON CHILDREN WITH DISABILITIES, SECTION ON DEVELOPMENTAL AND BEHAVIORAL PEDIATRICS. Identification, Evaluation, and Management of Children With Autism Spectrum Disorder. Pediatrics. 2020;145(1):e20193447. doi: 10.1542/peds.2019-3447.

33. Hosmer DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression. 3rd ed. Hoboken (NJ): Wiley; 2013.

34. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. Stat Med. 2019;38:2074–2102.