

Immortal time bias for life-long conditions in retrospective observational studies using electronic health records

Freya Tyrer (✉ fct2@le.ac.uk)

University of Leicester

Krishnan Bhaskaran

London School of Hygiene & Tropical Medicine

Mark J Rutherford

University of Leicester

Research Article

Keywords: bias (epidemiology), epidemiologic methods, immortal time bias, electronic health records, life expectancy, observational studies

Posted Date: November 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1060955/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Immortal time bias is common in observational studies but is typically described for pharmacoepidemiology studies where there is a delay between cohort entry and treatment initiation.

Methods

This study used the Clinical Practice Research Datalink (CPRD) and linked national mortality data in England from 2000–2019 to investigate immortal time bias for a specific life-long condition, intellectual disability. Life expectancy (Chiang's abridged life table approach) was compared for 33,867 exposed and 980,586 unexposed individuals aged 10+ years using five methods: (1) treating immortal time as observation time; (2) excluding time before date of first exposure diagnosis; (3) matching cohort entry to first exposure diagnosis; (4) excluding time before proxy date of entering first exposure diagnosis (by the physician); and (5) treating exposure as a time-dependent measure.

Results

When not considered in the design or analysis (Method 1), immortal time bias led to disproportionately high life expectancy for the exposed population during earlier calendar periods (additional years expected to live: 2000–2004: 65.6 [95% CI: 63.6,67.6]; 2005–2009: 59.9 [58.8,60.9]; 2010–2014: 58.0 [57.1,58.9]; 2015–2019: 58.2 [56.8,59.7]). Date of entry of diagnosis (Method 4) was unreliable in this CPRD cohort. The final methods (Method 2, 3 and 5) appeared to solve the main theoretical problem but residual bias may have remained.

Conclusions

We conclude that immortal time bias is a significant issue for studies of life-long conditions that use electronic health record data and requires careful consideration of how clinical diagnoses are entered onto electronic health record systems.

Background

Electronic health records are increasingly being used to conduct real-world observational studies to determine the association between a treatment or exposure and outcome. However, such studies are prone to a number of biases. In particular, immortal time bias is a recognised limitation of observational studies [1–3] that has come to the forefront in recent years owing to the increasing complexity of observational cohort designs [4–8]. This bias occurs where there is a period of time during follow up where an event or death cannot occur [4]. It is often discussed in the context of pharmacoepidemiology

where a delay between entering the study and being allocated a given therapy at baseline creates an ‘immortal’ period for the subject thereby creating an apparent advantage for the group that is given the therapy. In these types of study, immortal time bias can generally be reduced, if not eliminated completely, by adapting the analysis; for example, through application of prescription time-distribution matching (PTDM), time-dependent or sequential Cox approaches, or landmark analyses [9–11].

However, immortal time bias is not specific to pharmacoepidemiology studies and can represent a significant problem where the exposure is a life-long condition or disability. Whilst it may be difficult for some to consider a life-long condition to be an exposure measure, this is the best way to conceptualise factors associated with the outcome of interest for epidemiological purposes. Where the exposure is life-long, it may not be possible to completely control for the effects of immortal time bias in the design or analysis. We demonstrate this using an example of electronic health record data of life expectancy among patients with and without intellectual disabilities from the UK Clinical Practice Research Datalink (CPRD). Our findings may also have applicability to long-term conditions with delays in diagnosis or prolonged latency periods.

Methods

Source of data

For this example, we used the Clinical Practice Research Datalink (CPRD GOLD), linked (person-level) with hospital episode statistics (HES) and death registrations from the Office for National Statistics (approved study protocol number: 19_267RA3). The CPRD is an electronic health record research database of more than 11.3 million patients, broadly representative of the national population in terms of age, gender, and ethnicity [12], from general practice (GP) surgeries in the UK – of which approximately 75% in England consent to linkage to deaths data. The study followed the Reporting of studies Conducted using Observational Routinely-collected health Data (RECORD) checklist [13] (see supplementary **Table S1**).

Diagnostic codes used in this study are reported in supplementary **Table S2**. The initial extract from the CPRD has been described previously [14] and was based on the following inclusion criteria: registered at the GP surgery at any time between 1 Jan 2000 to 29 Sept 2019, and 10 years old or over to account for delays in reporting of diagnoses of intellectual disability in children [15]. An additional 23 patients with Angelman or Cockayne syndrome were added in August 2021 after an amendment to the original protocol (approved March 2020 but delayed during the COVID period). A random sample of people without intellectual disabilities was used for the comparison group with the same eligibility criteria (but without a diagnosis of intellectual disability). The initial extract included 33,867 people with intellectual disabilities (the exposed population) and a random sample of 980,586 people without intellectual disabilities (the unexposed population), although population sizes varied by the immortal time bias approach adopted. The baseline characteristics of the study populations under the five approaches are shown in supplementary **Table S3**.

Exposure/control definitions to handling immortal time bias under five different approaches

We present five approaches to defining cohort entry time when calculating life expectancy in people with and without intellectual disabilities and describe the impact that each approach has on life expectancy estimates in the context of immortal time bias. All methods involved changes in the handling of the exposed population. The second and third method also involved changes to the unexposed (control) population. The five methods compared are summarised in **Figure 1**.

Method 1: Treating immortal time as observation time

Method 1 involved applying no additional cohort entry criteria to either exposed or unexposed populations such that both populations are treated in the same way.

Method 2: Excluding immortal time before date of first exposure diagnosis

Method 2 involved adding date of intellectual disability diagnosis to the entry criteria such that date of entry into the cohort was set to the date of intellectual disability diagnosis if this was *after* the entry criteria defined for Method 1 (**Figure 1**). The comparison group patients entered at the date of registration/start of follow-up, as in Method 1. This approach has been described in pharmacoepidemiology studies in the context of excluding immortal time prior to treatment initiation in the treated group [10].

Method 3: Matching cohort entry to first exposure diagnosis

The third approach involved excluding immortal time, as in Method 2, but then matching the exposed individuals to unexposed individuals at a 1:10 ratio by date of cohort entry. This approach was designed to give a more balanced distribution of cohort entry date than Method 2 (in which entry dates for the comparison group tended to be earlier by design). The method is similar to the PTDM approach described for pharmacoepidemiology studies where dates of initiating therapy vary between treatment groups [10, 16, 17]. The difference between this and our approach is that PTDM involves differentiating between the 'never treated' and 'ever treated' groups such that cohort entry dates in the 'never treated' group are shifted to the date that the 'ever treated' group first started their treatment [10]. The approach, therefore, requires conditioning on the future and depends on length of the follow-up period since those in the 'never treated' group are allowed to move to the 'ever treated' group if the follow-up period is sufficiently long. Instead, our approach involved matching cohort entry dates against unexposed individuals, which included 'never exposed' individuals (i.e. without intellectual disabilities; PTDM approach) and 'ever exposed' individuals (i.e. people with intellectual disabilities *prior to* their first diagnosis).

Ten matches for each exposed individual were initially selected from the pool of unexposed individuals who entered the study within 150 days (to balance follow-up time) prior to the index date (i.e. date of cohort entry for exposed individuals). Where all 10 matches could not be found within this time constraint (n=24 patients), controls were selected from the pool of unexposed individuals who entered at any time

prior to the index date and were still at risk. Dates of cohort entry in the unexposed population were then updated to the index date.

Method 4: Excluding time before proxy date of entering first exposure diagnosis

The fourth approach involved incorporating the date that the exposed population's diagnosis was assumed to be entered by the physician onto the GP surgery's electronic health record system. This is not typically an issue with pharmacoepidemiology studies because the date of prescribing treatment is usually close to the date that the individual commences the treatment. However, for a life-long condition such as intellectual disabilities, physicians may choose to backdate the first exposure diagnosis to the patient's date of birth. The CPRD provides a linked variable ('system date') for each corresponding diagnosis date which may correspond to the date that the diagnosis was entered. However, this variable is also updated when a person transfers to the GP surgery, when there is an update in the GP software system used, or when the record is amended [18] which may lead to erroneous loss of person-years in the exposed group. To investigate this, date of entry was set to the date attached to the first intellectual disability diagnosis if it was later than the date of entry defined in Method 1 or Method 2 (by design; Figure 1). In this method, the unexposed patients entered at the date of registration/start of follow-up period as in Method 1.

Method 5: Treating exposure as a time-dependent measure

The final approach to handling immortal time bias involved treating the exposure as a time-dependent variable. This is perhaps the most common approach used in pharmacoepidemiology studies to control for immortal time bias and involves adapting the analysis so that individuals' exposure to a therapy is allowed to change during the follow-up period [10, 17]. It can be used to investigate the effect of individual therapies where more than one is under investigation [19] and is also advocated as a means of avoiding immortal time bias caused by temporal variability in the onset of certain conditions, such as the menopause [3]. In this study, the approach involved allowing individuals in the exposed population to contribute to the unexposed population until their first intellectual disability diagnosis whereupon they started contributing to the exposed population (**Figure 1**). The advantage of this method is that people with intellectual disabilities contribute person years in the same way as those who died before they had the opportunity to have an intellectual disability diagnosis.

Statistical Analyses

For the purposes of this work, data were split into the following calendar periods: 2000–2004; 2005–2009; 2010–2014; and 2015–2019. Date of entry into the cohort was defined as the latest date according to the person and practice's characteristics: 01 Jan 2000; date of registration with the GP practice; date the practice was defined as being up to standard (using the CPRD's own quality indicators); or date the individual turned 10 years old (to align with the eligibility criteria). Additional entry cohort criteria were

specified according to the approach used. Date of exit from each calendar period was calculated in the same way throughout as: date of death; date of end of calendar period; date of last practice update (latest 29 Sep 2019); or date of transfer out of practice, whichever was first. If the patient died after their date of exit from the cohort, they were censored on the date of exit.

To calculate life expectancy (additional life years expected to live) in both the exposed and unexposed populations, the Chiang's abridged life table approach [20–23] was used. This approach has been described in detail elsewhere [23], but briefly involves stratifying by exposed and unexposed status and constructing a table of probabilities that individuals will survive in a defined age interval conditional on surviving to the start of that age interval. The product of probabilities is then used to calculate survival to each age interval and life expectancy is estimated as the cumulative number of years lived using information from all subsequent age intervals divided by the population at the start of the given age interval. Five-year-period life expectancies were estimated to highlight the immortal bias findings, allowing that there may have been period and cohort effects during each five-year calendar period.

Confidence intervals for the derived life expectancies were calculated using the Chiang II approach as advocated by Eayres & Williams [24]. This involves adding a correction term to the original Chiang variance to account for the under-estimation of the 'true' variance at the last age interval by assuming variance is zero rather than basing the estimate on length of survival [25].

Results

Figure 2 summarises the life expectancy findings for the exposed population under the five methods. Each of the methods is described in more detail below.

Method 1: Treating Immortal Time As Observation Time

We can see from **Figure 2** that there is an apparent survival advantage in the first calendar period for the exposed population using Method 1. At age 10 years, for example, the estimated additional years expected to live in 2000–2004 was 65.6 (95% confidence interval [CI] 63.6,67.6) compared with 59.9 (58.8,60.9), 58.0 (57.1,58.9) and 58.2 (56.8,59.7) in the subsequent calendar periods. In the first calendar period and under Method 1, 38.5% of person-time ($n=21,506$ person years) in the cohort with intellectual disabilities was *before* the first intellectual disability diagnosis, compared with 23.4%, 9.1% and 4.8% in the second, third and fourth calendar periods respectively. Therefore, person-time was accrued for the exposed population during which a death could not occur.

Method 2: Excluding immortal time before date of first exposure diagnosis

Under Method 2, we can now see that the life expectancy advantages in the first calendar period are not as apparent. However, life expectancy remains slightly higher compared with the other calendar periods. This could be a real effect, but we speculate that some of the GP surgeries may have backdated

intellectual disability diagnoses to date of birth during some key periods in response to policy initiatives in England (see discussion). In the first calendar period, 14.5% of person time ($n=8,091$ person years) had a backdated intellectual diagnosis to year of birth. This compares with 16.2% and 17.0% in the second and third calendar period during which many of the policy initiatives occurred and 14.6% in the final calendar period. We are unable to determine whether these records were updated when the patient was registered with the practice (when immortal time bias would not present a problem) or at some point during the registration period (see **Method 4** for one approach to handling this).

There is also an additional problem that, by excluding immortal time from the observation period, we introduce time-related bias by forcing many of the exposed individuals to enter later than the unexposed population (**Figure 1**). Under Method 2, the median length of follow-up was shorter in the exposed population (4.6yrs vs 5.0 yrs; **Table S3**). The percentage of people with intellectual disabilities in the entire cohort (i.e. combining the cohort with and without intellectual disability) was also smaller in the first year of cohort entry period, increasing thereafter (see **Figure 3**). Bias introduced from this approach is believed to be negligible if the person-years in the exposed cohort is much smaller than the person-years in the unexposed cohort [10], but this does mean that a smaller sample of people with intellectual disabilities are investigated in earlier cohort periods. The next method attempts to evaluate this using matching techniques.

Method 3: Matching Cohort Entry To First Exposure Diagnosis

It can be seen that matching on cohort entry in addition to excluding immortal time before date of first intellectual disability diagnosis (Method 2) resulted in a more balanced distribution of people in the exposed population over the observation period (**Figure 3**). As expected, the life expectancy calculations for exposed individuals were identical to Method 2 (**Figure 2**). Life expectancy was also similar in the unexposed group (**Figure 4**), although person-time contribution was smaller because 65.5% ($n=641,916$) of individuals from this population were discarded from the analysis because they were not matched.

Method 4: Excluding time before proxy date of entering first exposure diagnosis

Method 4, implemented under the assumption that the CPRD's system date represented the date that the physician entered the intellectual disability diagnosis, produced the lowest life expectancy estimates for people with intellectual disabilities in the earlier calendar year periods (**Figure 2**). All calendar periods, except the last calendar period, showed a markedly lower life expectancy for people with intellectual disabilities than the other two methods. The graphs show that life expectancy estimates in the general population for this method increased only slightly over time (**Figure 4**) with life expectancy in people with intellectual disabilities showing more dramatic improvements (**Figure 2**).

However, this method led to a substantial loss in person-year contributions (median follow-up period 2.2yrs vs 5.0yrs) in the earlier cohort periods (**Figure 3**). During the first calendar period (2000–2004),

almost three-quarters of person-year contributions were lost compared with Method 1 (78.4%, $n=43,787$ yrs) and two-thirds compared with Method 2 (64.9%, $n=22,281$ yrs). Person-years lost gradually lowered for the subsequent calendar periods (64.1% and 53.3%, respectively, for 2005–2009, 41.2% and 35.3% for 2010–2014; 24.0% and 20.2% for 2015–2019). We can see, therefore, that people with a good prognosis (i.e. surviving long enough to have a system update at some later point) may have systematically been removed. The supplementary **Figure S1** shows the distribution of system dates that were later than the first intellectual disability diagnosis. There were clusters of system dates linked to intellectual disability diagnoses at specific times that appeared to correspond to software system updates, particularly as some ($n=623$ individuals) occurred after the individual had died or left the practice (see **Figure S1**). Moreover, the use of system date as a proxy date for entering the cohort is not advocated for CPRD data because it is not sufficiently specific [26].

Method 5: Treating Exposure As A Time-dependent Measure

The final approach, treating the exposure as a time-dependent variable, showed very similar results to Method 2 and 3 for the exposed population (**Figure 2**) and marginal differences to the life expectancy calculations in the unexposed population (**Figure 4**) because the additional person-year contribution from the exposed population was relatively small. As before, this method has similar disadvantages to Method 2 in that we force exposed individuals to enter later than the unexposed population. In fact, the unexposed group contribute partially more to the earlier time periods than for Method 2.

Discussion

Through the use of electronic health records data, we demonstrate that immortal time presents a significant problem for time-to-event analyses where one or more of the exposures is a life-long condition or disability. Treating immortal time as exposure time (Method 1) led to an over-estimation of life expectancy advantages in the exposed population. Even when immortal time was excluded or treated as unexposed time (Methods 2, 3 and 5), some residual immortal time bias may have remained where diagnoses had been backdated to date of birth. Using proxy date of physician's entry of exposure diagnosis (Method 4) resulted in a substantial loss of person-time and subjects, and did not appear to be used in a consistent way in our data source. Our findings highlight that interpretation is key for any study where the exposure can occur after the start of the follow-up period and consideration of immortal time bias is needed to avoid drawing incorrect conclusions.

To our knowledge, this is the first time that the issue of immortal time bias has been studied in detail for life-long conditions or disabilities, although there is a wealth of literature that discusses immortal time bias in the context of pharmacoepidemiology studies. This literature largely corresponds with Methods 3 and 5 in our study by recommending the control of immortal time bias through PTDM or time-dependent approaches that allow exposure status to vary [9, 10, 17]. Time-dependent analyses have also been

recommended to control for the effects of immortal time bias where time of onset of certain health conditions varies, such as the menopause [3]. The reason that the issue of immortal time bias for life-long conditions/disabilities has not been considered before may be that it seems conceptually inappropriate to consider someone with a life-long condition to be 'disease-free' for any period of observation. However, this is likely to be the best solution for electronic health records that do not follow up individuals from birth.

The magnitude of immortal time bias is reported to be related to mean interval between date of cohort entry and date of (recording of) exposure, proportion of exposed study participants, and length of study follow-up [27]. The prevalence of intellectual disability, as diagnosed in primary care, is approximately 0.5% [28] but the current study population had a larger proportion of exposed individuals (~3%) because only a proportion of the unexposed population was selected for comparison. The cohort period of almost 20 years also increased the likely bias introduced by immortal time. Similarly, the choice of intellectual disability as an example may have led to substantially more immortal time bias than some other conditions or disabilities. Intellectual disability itself does not require treatment so, for administrative purposes, may not need to be entered onto GP systems if it is already known or reported in the patients' notes. This changed in 2004, with the introduction of the Quality Outcomes Framework (QOF) and incentives to report long-term conditions including intellectual disabilities [29, 30], closely followed by policy drives to maintain practice-level intellectual disability registers in 2006 (adults) and 2014 (children) [31–33], and annual intellectual disability health checks in 2008 (adults) and 2014 (14–17 year olds) [31, 33]. We cannot identify another condition or disability where a policy drive has been so influential in changing practice in primary care. However, increased awareness is known to 'artificially' increase incidence of certain conditions over time, such as autism and coeliac disease [34, 35]. Our findings may also have applications for conditions where there is a delay between onset and diagnosis, such as Crohn's disease or rare diseases [36, 37].

Finally, we have shown that attempts to control for immortal time bias in the design or analysis stage does not guarantee unbiased results. However, treating immortal time as observation time, thereby ignoring immortal time bias completely, is not recommended as we have shown that this will lead to spurious results owing to the 'misclassification of immortal time' [10]. Methods that start follow-up for exposed individuals at diagnosis (Methods 2, 3 and 5) all produce fairly similar findings and appear to solve the main theoretical problem, so any of these approaches could be adopted without a reference standard on which to compare. Although these methods all showed similar results for this study, Methods 3 and 5 have conceptual advantages over Method 2 because they do not involve conditioning on the future [38]. In other words, they allow exposed individuals to contribute to the unexposed population prior to their first diagnosis when they are still "at risk". Incorporating date of assumed entry of exposure diagnosis by the physician (Method 4) may have potential in some data sources but we would not recommend this as a primary analysis for studies that use the CPRD owing to the apparent inconsistent use of the system date field that we used to capture date of record entry and substantial loss to person-time and subjects. Studies of electronic health records where date of entry of diagnosis is more reliably recorded may be considered for this purpose. This approach may also be valuable for conducting

sensitivity analyses where immortal time bias is not perceived to be adequately controlled. Finally, we recommend adding an assessment of immortal time bias as a key component of critical appraisal tools for electronic health record studies.

Declarations

Ethics approval

This study used an established research database (CPRD) under a multi-study licence agreement with the UK Medicines and Healthcare products Regulatory Agency. The CPRD has ethical approval as a research database for anonymised data (East Midlands-Derby, REC reference number 05/MRE04/87). The protocol for this research was approved by the CPRD's Independent Scientific Advisory Committee (ISAC) (protocol number: 19_267). All methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication

Not applicable

Availability of data and materials

Data for this study were obtained from the Clinical Practice Research Datalink (CPRD), provided by the UK Medicines and Healthcare products Regulatory Agency. The authors' licence for using these data does not allow sharing of raw data with third parties. Information about access to CPRD data is available here: <https://www.cprd.com/research-applications>. Researchers should contact the ISAC Secretariat at isac@cprd.com for further details.

Competing interests

None declared

Funding

This research was funded from a Baily Thomas Doctoral Fellowship award (TRUST/VC/AC/SG/5366-8393) and the Wellcome Trust (107731/Z/15/Z and 220283/Z/20/Z). For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

FT led on the planning, conduct, design, reporting of the work, funding acquisition, data acquisition and analysis, and interpretation of data for this work. KB contributed to the planning, conduct, data analysis, interpretation of data and reporting. MR contributed to the planning, conduct, design, reporting of the work, funding acquisition, data acquisition and analysis, and interpretation of data.

Acknowledgments

We gratefully acknowledge Rachael Williams, Observational Research Manager at the Clinical Practice Research Datalink (CPRD) for her valuable input and advice. We also acknowledge the Talk & Listen group (and facilitation by Pauline Ndigirwa), the Charnwood Action group, Amy Stanway, Sarah Stanyer, and Kate Dolan. The study is based in part on data from the CPRD GOLD database obtained under licence from the UK Medicines and Healthcare products Regulatory Agency. The provision of CPRD and linked data was through Leicester Real World Evidence (LRWE) Unit, which is funded by University of Leicester, National Institute for Health Research (NIHR) Applied Research Collaboration (ARC) East Midlands and Leicester NIHR Biomedical Research Centre. The interpretation and conclusions contained in this article are those of the authors alone and not necessarily those of the LRWE Unit, the NHS, the NIHR, or the Department of Health and Social Care.

References

1. Farmer R, Mathur R, Bhaskaran K, Eastwood SV, Chaturvedi N, Smeeth L: **Promises and pitfalls of electronic health record analysis.** *Diabetologia* 2018, **61**(6):1241-1248.
2. Suissa S, Azoulay L: **Metformin and the risk of cancer: time-related biases in observational studies.** *Diabetes Care* 2012, **35**(12):2665-2673.
3. Yadav K, Lewis RJ: **Immortal time bias in observational studies.** *Journal of the American Medical Association* 2021, **325**(7):686-687.
4. Lévesque LE, Hanley JA, Kezouh A, Suissa S: **Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes.** *British Medical Journal* 2010, **340**:b5087.
5. Suissa S: **Immortal time bias in observational studies of drug effects.** *Pharmacoepidemiology & Drug Safety* 2007, **16**(3):241-249.
6. Austin PC, Platt RW: **Survivor treatment bias, treatment selection bias, and propensity scores in observational research.** *Journal of Clinical Epidemiology* 2010, **632**:136–138.
7. Kim YG, Park DG, Moon SY, Jeon JY, Kim HJ, Kim DJ, Lee KW, Han SJ: **Hypoglycemia and dementia risk in older patients with type 2 diabetes mellitus: A propensity-score matched analysis of a population-based cohort study.** *Diabetes & Metabolism* 2020, **44**(1):125-133.
8. Zaccardi F, Jacquot E, Cortese V, Tyrer F, Seidu S, Davies MJ, Khunti K: **Comparative effectiveness of gliclazide modified release versus sitagliptin as second-line treatment after metformin monotherapy in patients with uncontrolled type 2 diabetes.** *Diabetes, Obesity & Metabolism* 2020.
9. Jones M, Fowler R: **Immortal time bias in observational studies of time-to-event outcomes.** *Journal of Critical Care* 2016, **36**:195-199.

10. Karim ME, Gustafson P, Petkau J, Tremlett H, The Long-Term Benefits and Adverse Effects of Beta-Interferon for Multiple Sclerosis Study Group: **Comparison of statistical approaches for dealing with immortal time bias in drug effectiveness studies**. *American Journal of Epidemiology* 2016, **184**(4):325-335.
11. Mi X, Hammill BG, Curtis LH, Lai EC, Setoguchi S: **Use of the landmark method to address immortal person-time bias in comparative effectiveness research: a simulation study**. *Statistics in Medicine* 2016, **35**(26):4824-4836.
12. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, Smeeth L: **Data resource profile: Clinical Practice Research Datalink (CPRD)**. *International Journal of Epidemiology* 2015, **44**(3):827-836.
13. Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM: **The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement**. *PLoS Medicine* 2015, **12**(10):e1001885.
14. Tyrer F, Morriss R, Kiani R, Gangadharan SK, Rutherford MJ: **Mortality disparities and deprivation among people with intellectual disabilities in England: 2000–2019**. *Journal of Epidemiology & Community Health* 2021, doi: [10.1136/jech-2021-216798](https://doi.org/10.1136/jech-2021-216798).
15. Florio T, Trollor J: **Mortality among a cohort of persons with an intellectual disability in New South Wales, Australia**. *Journal of Applied Research in Intellectual Disabilities* 2015, **28**(5):383-393.
16. Pariente A, Fourrier-Réglat A, Ducruet T, Farrington P, Béland SG, Dartigues JF, Moore N, Moride Y: **Antipsychotic use and myocardial infarction in older patients with treated dementia**. *Archives of Internal Medicine* 2012, **172**(8):648-653.
17. Dekkers OM, Groenwold RHH: **When observational studies can give wrong answers: the potential of immortal time bias**. *European Journal of Endocrinology* 2021, **184**(1):E1-e4.
18. Leite A, Andrews NJ, Thomas SL: **Assessing recording delays in general practice records to inform near real-time vaccine safety surveillance using the Clinical Practice Research Datalink (CPRD)**. *Pharmacoepidemiology & Drug Safety* 2017, **26**(4):437-445.
19. Morriss R, Tyrer F, Zaccardi F, Khunti K: **Safety of antidepressants in a primary care cohort of adults with obesity and depression**. *PLoS ONE* 2021, **16**:e0245722.
20. Chiang CL: **Variance and covariance of life table functions estimated from a sample of deaths**. In. Washington: National Center for Health Statistics 1967: 1-7.
21. Chiang CL: **Life table and mortality analysis**. In. Geneva: World Health Organization 1979: 1-399.

22. Chiang CL: **The life table and its construction**. In: *Introduction to stochastic processes in Biostatistics*. edn. New York: John Wiley & Sons 1968: 189–214.
23. Chiang CL: **The life table and its applications**. Malabar: Robert E Krieger Publishing Company 1984.
24. Eayres D, Williams, E.S. : **Evaluation of methodologies for small area life expectancy estimation**. *Journal of Epidemiology & Community Health* 2004, **58**:243–249.
25. Lo E, Vatnik D, Benedetti A, Bourbeau R: **Variance models of the last age interval and their impact on life expectancy at subnational scales**. *Demographic Research* 2016, **35**:399-454.
26. Williams R: **Personal communication [December 2020]**. In.; 2020.
27. Harding BN, Weiss NS: **Point: immortal time bias - what are the determinants of its magnitude?** *American Journal of Epidemiology* 2019, **188**(6):1013-1015.
28. Glover G, Williams R, Heslop P, Oyinlola J, Grey J: **Mortality in people with intellectual disabilities in England**. *Journal of Intellectual Disability Research* 2017, **61**(1):62-74.
29. Lester H, Campbell S: **Developing Quality and Outcomes Framework (QOF) indicators and the concept of 'QOFability'**. *Quality in Primary Care* 2010, **18**(2):103-109.
30. Roland M, Guthrie B: **Quality and Outcomes Framework: what have we learnt?** *British Medical Journal* 2016, **354**:i4060.
31. British Medical Association & NHS Employers: **Clinical directed enhanced services (DES) guidance for GMS contract**. In. London: Department of Health 2008.
32. British Medical Association/NHS England/NHS Employers: **2014/15 General Medical Services (GMS) Contract Quality and Outcomes Framework (QOF). Guidance for GMS Contract 2014/15. NHS England Gateway reference: 01264**. In. England: British Medical Association, NHS England, NHS Employers. Available from: http://www.mysurgerywebsite.co.uk/website/IGP367/files/Guidance_and_audit_requirements14-15contract.pdf (Accessed 19 Jan 2021) 2014.
33. General Practitioners Committee/NHS England/NHS Employees: **General Medical Services Contract 2014-2015. Guidance and audit requirements: NHS England Gateway reference: 01347**. In. England: General Practitioners Committee, NHS England & NHS Employees. Available from: http://www.mysurgerywebsite.co.uk/website/IGP367/files/Guidance_and_audit_requirements14-15contract.pdf (Accessed 19 Jan 2021) 2014.
34. Taylor B, Jick H, MacLaughlin D: **Prevalence and incidence rates of autism in the UK: time trend from 2004–2010 in children aged 8 years**. *BMJ Open* 2013, **3**(10):e003219.

35. King JA, Jeong J, Underwood FE, Quan J, Panaccione N, Windsor JW, Coward S, deBruyn J, Ronksley PE, Shaheen AA *et al*: **Incidence of Celiac disease is increasing over time: A systematic review and meta-analysis.** *American Journal of Gastroenterology* 2020, **115**(4):507-525.
36. Blackwell J, Saxena S, Jayasooriya N, Bottle A, Petersen I, Hotopf M, Alexakis C, Pollok RC, group P-Is: **Prevalence and duration of gastrointestinal symptoms before diagnosis of inflammatory bowel disease and predictors of timely specialist review: A population-based study.** *Journal of Crohn's and Colitis* 2020.
37. Evans WR, Rafi I: **Rare diseases in general practice: recognising the zebras among the horses.** *Br J Gen Pract* 2016, **66**(652):550-551.
38. Lund JL, Horváth-Puhó E, Komjáthiné Szépligeti S, Sørensen HT, Pedersen L, Ehrenstein V, Stürmer T: **Conditioning on future exposure to define study cohorts can induce bias: the case of low-dose acetylsalicylic acid and risk of major bleeding.** *Clin Epidemiol* 2017, **9**:611-626.

Figures



Figure 1

Diagram of exposed person-time under five methods for studies of life-long conditions using electronic health record data

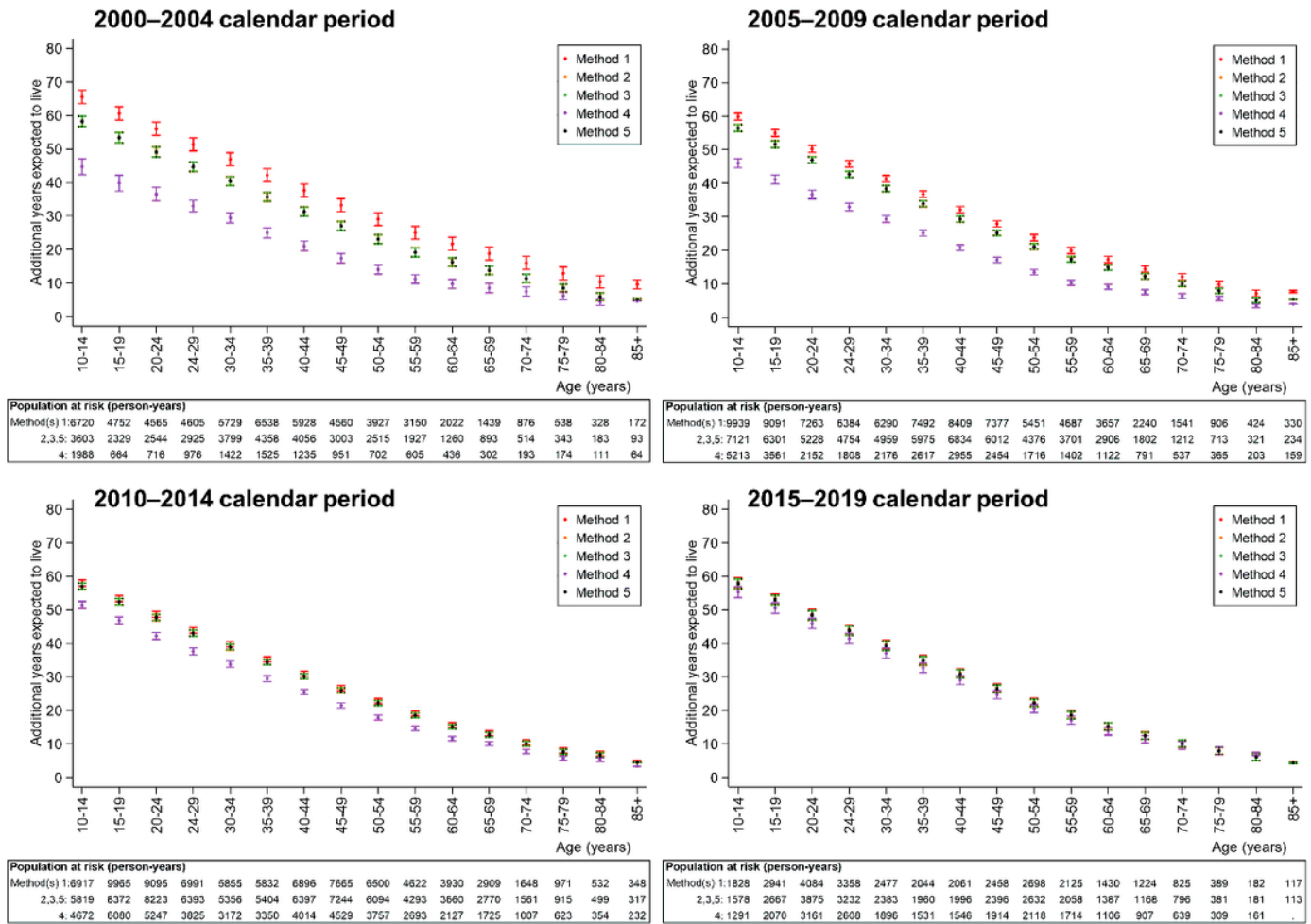
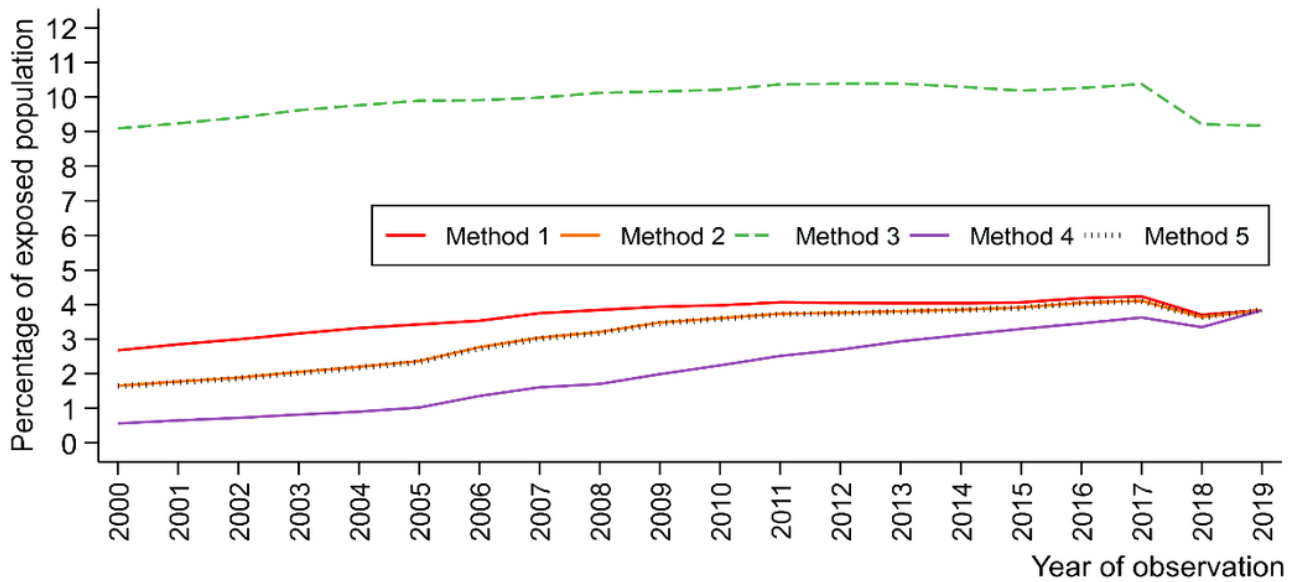


Figure 2

Life expectancy by calendar period in people with intellectual disabilities under the five methods to handle immortal time bias



Population at risk (person-years)										
EXPOSED										
Method(s) 1:	9014	12303	14615	16354	18358	19126	18240	15641	9197	4814
2,3,5:	5503	7660	9569	12712	15212	17263	16901	14888	8888	4724
4:	1850	2909	3875	6142	7960	10586	11981	11965	7533	4331
UNEXPOSED										
1,2,4:	327798	398812	425728	446916	459413	461634	432322	371710	210507	125137
3:	55030	73801	88447	115607	135065	151816	145785	129661	77742	46545
5:	331536	403807	431248	452392	463331	464248	434218	372964	211092	125393

Figure 3

Percentage of exposed individuals by year of observation a, b. a Method 4 can contain more than one individual where first intellectual disability diagnosis date is greater than the date of entry (e.g. a person entering the cohort in 2000 but diagnosed first with intellectual disability in 2006 enters the cohort without intellectual disabilities in 2000 and enters again with intellectual disabilities in 2006). b Please note that, as a random sample of the general population without intellectual disabilities for the comparison group, this graph cannot be interpreted as representing prevalence of intellectual disability.

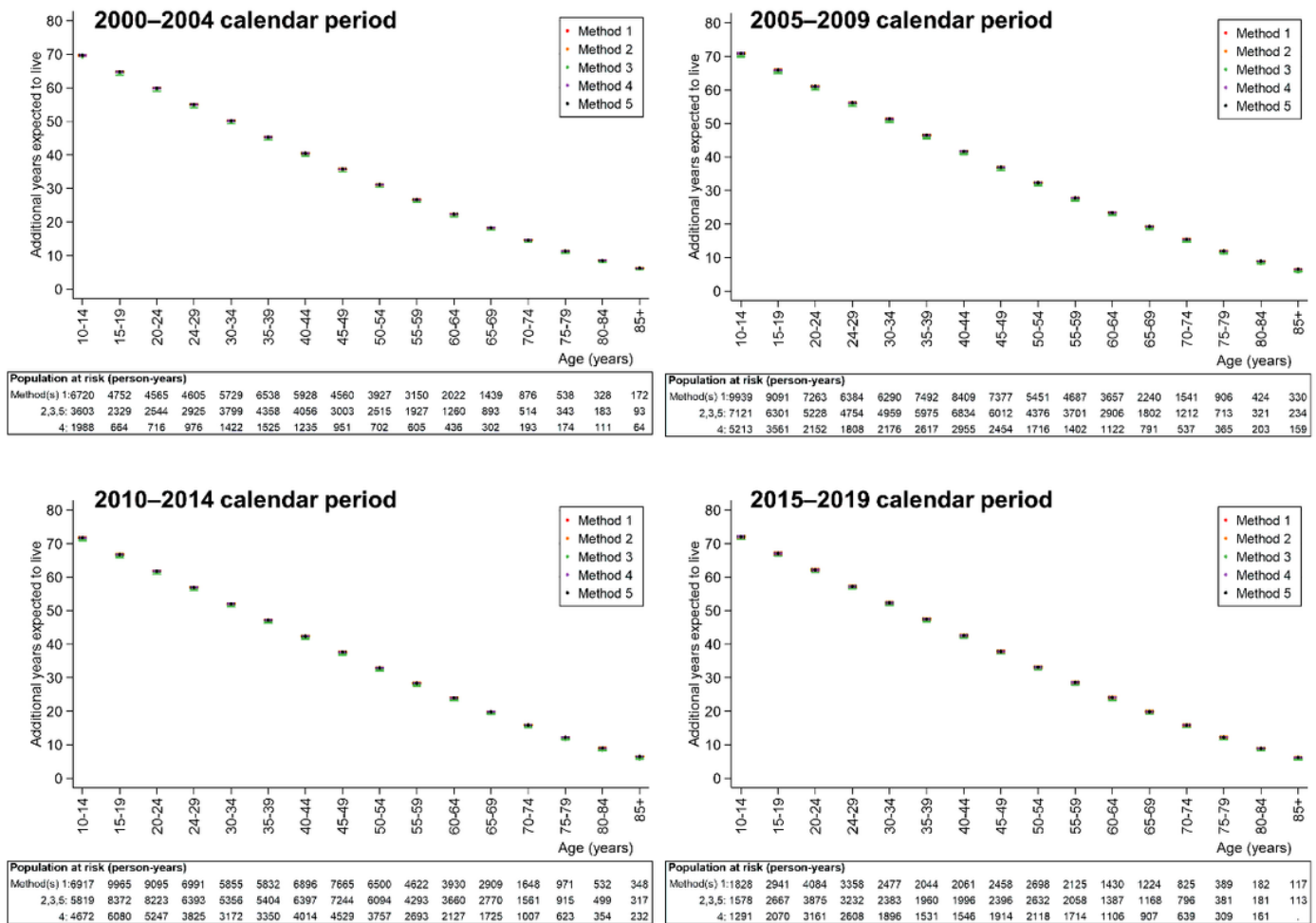


Figure 4

Life expectancy by calendar period in people without intellectual disabilities under the five methods to handling immortal time bias

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TyrrerSupplementaryMaterialFINAL112021.docx](#)