HHS Public Access

Author manuscript

Int J Stat Med Res. Author manuscript; available in PMC 2016 April 25.

Published in final edited form as:

Int J Stat Med Res. 2016 ; 5: 8–16. doi:10.6000/1929-6029.2016.05.01.2.

Use of Self-Matching to Control for Stable Patient Characteristics While Addressing Time-Varying Confounding on Treatment Effect: A Case Study of Older Intensive Care Patients

Ling Han[1,*], **M.A. Pisani**[1,2], **K.L.B. Araujo**[1], and **Heather G. Allore**[1,3]

[1]Department of Internal Medicine, Program on Aging, Yale School of Medicine, New Haven, CT, USA

[2]Department of Internal Medicine, Section of Pulmonary, Critical Care and Sleep Medicine, Yale School of Medicine, New Haven, CT, USA

[3]Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA

Abstract

Exposure-crossover design offers a non-experimental option to control for stable baseline confounding through self-matching while examining causal effect of an exposure on an acute outcome. This study extends this approach to longitudinal data with repeated measures of exposure and outcome using data from a cohort of 340 older medical patients in an intensive care unit (ICU). The analytic sample included 92 patients who received ≥1 dose of haloperidol, an antipsychotic medication often used for patients with delirium. Exposure-crossover design was implemented by sampling the 3-day time segments prior (*Induction)* and posterior (*Subsequent)* to each *treatment episode* of receiving haloperidol. In the full cohort, there was a trend of increasing delirium severity scores (Mean±SD: 4.4±1.7) over the course of the ICU stay. After exposure-crossover sampling, the delirium severity score decreased from the *Induction* (4.9) to the *Subsequent* (4.1) intervals, with the *treatment episode* falling in-between (4.5). Based on a GEE Poisson model accounting for self-matching and within-subject correlation, the unadjusted mean delirium severity scores was −0.55 (95% CI: −1.10, −0.01) points lower for the *Subsequent* than the *Induction* intervals. The association diminished by 32% (−0.38, 95%CI: −0.99, 0.24) after adjusting only for ICU confounding, while being slightly increased by 7% (−0.60, 95%CI: −1.15, −0.04) when adjusting only for baseline characteristics. These results suggest that longitudinal exposure-crossover design is feasible and capable of partially removing stable baseline confounding through self-matching. Loss of power due to eliminating treatment-irrelevant person-time and uncertainty around allocating person-time to comparison intervals remain methodological challenges.

[*]Address correspondence to this author at the Yale Program on Aging, 300 George St Suite 775, New Haven CT 06511, USA; Tel: 203-737-2162; Fax: 203-785-4823; ling.han@yale.edu.

Keywords

Exposure-crossover design; self-matching; confounding; causal effects; generalized estimating equation

INTRODUCTION

Self-matching offers an attractive study design option alternative to randomization and peer-matching for observational studies of treatment effect, i.e., control for selection bias and confounding due to stable, patient-specific risk factors, such as genetic predisposition, personality traits and education attainment [1, 2]. Such stable confounding is inherent to observational data and often difficult to address in the analytic stage [1, 2]. The self-matching principle was first applied in epidemiological literature through case-crossover design [1], which was devised to examine the *transient* exposure effects on an *acute* outcome through comparing a designated "case" period and one or more "control" periods. A variant of case-crossover approach, called "exposure-crossover", was proposed [3], which shares several key features as its precursor, such as using each subject as their own control and comparing the outcome risks during an assumed effect period and a control period, yet anchors the analyses on the time of exposure instead of the outcome (or case) [3].

Since its introduction, at least two population-based studies have used the exposure-crossover approach to address the risk of an adverse outcome in an administrative database [4, 5]. The two studies defined a 1-year post-exposure period (or *subsequent interval*) to detect effects of fibromyalgia diagnosis on motor vehicle crashes or a medication injection on thromboembolism risk, in reference to a broader 3-year control period (or *baseline interval),* and found a significant detrimental association. However, whether this novel approach can be applied to longitudinal studies with repeated measures of exposure and outcome, especially in the quasi-experimental context, has not been explored in the literature.

This study attempts to extend the exposure-crossover approach to longitudinal data with multiple episodes of medication treatment over time. We illustrate our approach using a cohort of elderly patients receiving intensive care who have multiple comorbidities and are simultaneously receiving several medications (or polypharmacy). The scientific question behind this exercise is whether the administration of haloperidol, an antipsychotic medication commonly used to treat delirious patients, reduces the severity of delirium, an acute confusional state. Previous studies including sparse clinical trials have been insufficient and sometimes conflicting regarding the effectiveness of haloperidol in treating delirium, calling for novel observational studies to fill in the knowledge gap [6, 7].

METHODS

Prototype of Exposure-Crossover Design

In the seminal paper by Redelmeier [3], the exposure-crossover design involves 3 major steps of data reorganization. First, establish a time zero based on the time of exposure for each subject. Second, follow each subject for outcome experience both backward (pre-

treatment) and forward (post-treatment) from the defined time zero. Third, collapse the entire timeline of the study period into 3 sequential intervals, called the *baseline, induction* and *subsequent* intervals, respectively. In the analyses, the "causal" effect of the exposure is estimated by comparing the *subsequent (*serving to detect and quantify post-exposure outcomes) and the *baseline* (serving to detect and quantify long-term temporal trends prior to the exposure) intervals, while excluding the *induction* interval (reflecting nuisance related to reverse causality, confounding by indication, or other biases) [3]. To ensure a fair and efficient comparison, each interval was divided into time segments with uniform duration, typically by calendar year or 13 segments of 28-days [3].

Adapt Exposure-Crossover Design to Repeated Measure Data

To examine the "causal" effect of repeated haloperidol treatments among older ICU patients with multi-morbidities and polypharmacy regimen, we adapted the exposure-crossover design in the following aspects:

1. **Define *Treatment Episodes* of Consecutive Haloperidol Doses as Time Zero**

   In previous studies of exposure-crossover design, exposure was typically assumed to occur at a single time point [3–5] and embedded into the *induction* interval as a nuisance. In our sample, however, patients may receive haloperidol treatment over several consecutive days. To make an optimal use of such repeated exposure data while retaining its essential feature of self-matching, we first identified all the time periods of 1 or more days in which a patient was given a haloperidol dose every day as a *treatment episode*. The first and last day of these *treatment episodes,* respectively, were then used as a *time zero* to define the *induction* (pre-treatment) and *subsequent* (post-treatment) intervals. However, the outcome assessments on delirium severity during the treatment episodes may intervene with serial administrations of haloperidol doses and the (residual) effects of the medication and the delirium pathology are difficult to disentangle from each other. To avoid causal ambiguity, we considered the *treatment episodes* nuisance parameter and only evaluated the delirium severity in these time periods in descriptive and sensitivity analyses.

2. **Construct Comparison Intervals Based on Biologically Plausible Exposure Window**

   After identifying all haloperidol *treatment episodes*, the next step was to sample time segments to construct comparison intervals, i.e., the *baseline*, *induction* and *subsequent* intervals, relative to the defined time zero. To appropriately assign person-time segments to each interval while avoiding "pragmatics" and arbitraracy [3], we first need to define a biologically plausible exposure window for detecting the "causal" effect of haloperidol treatment on the delirium symptoms. Based on the fact that delirium symptoms typically arise in hours to days following exposure to some causal risk factors and that the half-life for haloperidol ranges from 14 to 24 hours [8], we considered the 3 days immediately following a haloperidol dose be the most plausible exposure window. Accordingly, we assigned a maximum 3 days and a minimum 1-day after each *treatment episode* to the *subsequent* intervals. We

then tried to allocate an equal number of maximum 3 days preceding each *treatment episode* to the *induction* intervals as the comparison period. When number of person-days between 2 adjacent *treatment episodes* is <6, we assign the days to *subsequent* interval first, under the minimum 1-day exposure window assumption. If available person-days for allocation is>6, the extra days after assigning the 3-days to each of the 2 intervals were excluded.

3. **Omit Defining a *Baseline Interval* as a Comparison Period**

   While the prototypical exposure-crossover design requires defining a *baseline* interval to capture the ongoing outcome risk unrelated to exposure as a comparison period [3–5], we chose not to do so for both theoretical (retaining the self-matching principle) and practical (lack of data prior to ICU admission) reasons. First and foremost, the scientific question prompting this self-matching design is whether the change in delirium severity is caused by haloperidol treatment, or by other "triggering" causes or temporal confounding proximal to the time of treatment [1,2,9]. Therefore, the most relevant time periods to detect such a change would be the time periods immediately preceding (or *induction* to) and following (or *subsequent* to) the exposure, rather than a distant period farther before the exposure. In addition, older patients entering the ICU are known to have an increased risk of delirium. Such ongoing risk may result from cumulative effects of genetic predisposition, long-term exposure to environmental factors, as well as the natural aging process, and is not expected to change dramatically over a short period prior to ICU admission (and thus, presumably are "stable") [1–3]. Finally, the salient point of the exposure-crossover design is its built-in self-matching capacity, whereby the stable baseline confounding, measured and unmeasured, is supposedly automatically removed. It would be redundant or counterintuitive to reiterate an explicit baseline period as a comparison or reference period. Therefore, instead of defining a baseline interval, we considered adjusting several patient characteristics at ICU admission in the analyses as proxy to stable baseline confounding.

A schematic illustration of this longitudinal version of exposure-crossover design is presented in Figure 1.

Data Source and Analytic Sample

To illustrate our approach, we used the Evaluation of Psychoactive Medications in the Intensive Care Unit (EPIC) which consists of 304 consecutive patients 60 years or older admitted to the medical intensive care units (ICU) of the Yale-New Haven Hospital, Connecticut, between September 5, 2002 and September 30, 2004 [7, 10]. Patients were assessed for delirium with the Confusion Assessment Method for the ICU (CAM-ICU) by a trained research nurse on a daily basis [11]. The dose, route and time of administration of all medications were tracked through 3 nursing shifts (day, evening and night). Other data collected included baseline demographics (e.g., age, sex, race), clinical characteristics (e.g., admitting diagnosis, dementia, and laboratory tests) and ICU interventions (e.g., mechanical ventilation), as described previously [7, 10]. During the ICU stay (mean duration ±SD:

7.2±7.9 days), 234 patients developed delirium, of whom 93 received at least one dose of haloperidol, the most common pharmacological treatment for delirium in clinical practice [6, 7]. After excluding one patient who received only a single dose at last day in ICU, we included 92 haloperidol users in the analytic sample.

Statistical Analyses

Study Variables—To derive a quantitative measure of delirium severity as study outcome, each of the 4 diagnostic features of the CAM-ICU, i.e., 1) acute onset or symptom fluctuation; 2) inattention; 3) disorganized thinking; and 4) altered level of consciousness, was assigned a score of 0 (absent), 1 (mild) or 2 (severe/marked), following a validated CAM-S algorism using the same 4 features [12]. The summary delirium severity score ranges from 0 to 7, with 7 indicating most severe symptoms.

The haloperidol treatment each patient received during ICU stay was measured as cumulative daily does in milligrams, and grouped into one or more discrete *treatment episode,* as defined above.

Several patient characteristics were used to capture stable baseline confounding, including age (in year), Acute Physiology and Chronic Health Evaluation (APACHE II) score, admission diagnoses (4 categories), dementia (yes vs no) and antipsychotic use at admission (yes vs no). ICU confounding was represented by a categorical variable denoting the time spent on mechanical ventilation during ICU stay (0, 1–80%, ≥80%) and cumulative anticholinergic daily drug burden as a time-varying continuous variable. The later was a standardized daily drug dose across 12 anticholinergic medications, a class of common geriatric medications known to cause delirium [13, 14].

Graphical Representation and Descriptive Analyses

We used line and bar-char plots to explore the distribution of delirium severity over the course of ICU stay in the original data and after applying exposure-crossover sampling [3]. We used mean and standard deviation for continuous variables, and frequency and percentages for categorical variables to summarize the population characteristics and exposure crossover design features.

Generalized Estimating Equation Modeling—We used a generalized estimating equation Poisson model [15] to examine the “causal” effects of haloperidol treatment by comparing the average delirium severity scores for the *subsequent* and the *induction* intervals while ignoring the *treatment episodes* to avoid potential reverse causality. We estimated the mean difference on the delirium severity scores between the two intervals through identity link function, with robust variance estimator for 95% confidence interval calculation. Self-matching was accounted for by treating each pair of comparison intervals (i.e., *induction* and *subsequent*) as a cluster. As patients could contribute multiple exposure-crossover intervals (and *treatment episodes*) to the analyses, we used an unstructured covariance to account for each persons’ unique correlation pattern. We systematically adjusted for the baseline characteristics at ICU admission and precipitating confounding during ICU stay, and jointly, and used the percent change in the parameter estimates from

the unadjusted model to each adjusted model to quantify the capacity of the exposure-crossover design in achieving self-matching.

We performed sensitivity analyses on the fully adjusted GEE model. First, we refit the model by including the *treatment episodes* as an explicit (nuisance) parameter, to test potential pragmatics and person-time selection bias [1, 2, 16]. Second, we restricted the analyses to a maximal 1-day duration for both *induction* and *subsequent* intervals, to examine the impact of unequal length assignment to the intervals. Third, we restricted the analyses to the first *induction* and the last *subsequent* intervals, to remove potential reverse causality bias due to residual effects of most recent haloperidol dose or delirium symptoms over successive intervals. Finally, we refit the model by ignoring the self-matching mechanism, to verify the notion of potential underestimation bias [1].

All statistical analyses were performed using SAS software version 9.4 (SAS Institute, Cary NC 2010), with a two-sided $\alpha = 0.05$ for statistical significance.

RESULTS

The characteristics of the 92 older patients included in the study sample are summarized in Table 1, which are largely comparable to the full cohort except for increased baseline cognitive impairment (IQCODE>3.3), less admissions for gastrointestinal hemorrhage and more time mechanically ventilated.

The 92 patients contributed a total of 825 person-days to the exposure-crossover sample. Of these, 398 person-days received ≥1 haloperidol dose at 6.6±10.8 mg per day and constituted 130 *treatment episodes*, with an average contribution of 1.4±0.9 episodes per person (range: 1–6). The delirium severity score on 251 person-days was 4.4±1.7 (range: 0–7).

During the ICU stay, the observed delirium severity scores among the full cohort appeared to increase gradually from admission to day 28, as shown in Figure 2.

After trimming the person-time data unrelated to the exposure-crossover design, the monotonic trend of delirium severity over time in disappeared, as shown in Figure 3.

Figure 4 presents the average delirium severity scores for the 3-day before (induction interval) and 3-day after (subsequent interval) the haloperidol treatment episode, aggregated across person-days in the exposure-crossover sample. The mean delirium severity scores were highest during the induction interval (4.9, 95% CI: 4.4, 5.4) and lowest during the subsequent interval (4.1, 95% CI: 3.8, 4.5), with the treatment episodes coming in-between (4.5, 95% CI: 4.1, 4.8).

As summarized in Table 2, there was a modest reduction in the delirium severity score from the induction to the subsequent intervals, with an unadjusted mean difference of -0.55 (95% CI: −1.10, −0.01) points. The association changed by 7% after adjusting for baseline characteristics (−0.60, 95% CI: −1.15, −0.04); yet diminished by 32% after adjusting for ICU covariates (−0.38, 95% CI: −0.99, 0.24). Adjusting for both baseline and ICU covariates did not alter the point estimate (−0.38, 95% CI: −1.01, 0.25).

Sensitivity analyses derived consistent results, with a statistically insignificant reduction of delirium severity score following the treatment episode (Table 3). In addition, the two extreme scenarios provided some support for the necessity of ensuring temporal unambiguity (model 3) [1, 2] and of accounting for self-matching in the analyses (model 4) [1]. The great variation of the point estimates may reflect the dramatically reduced sample size or potential over-parameterization due to estimating unstructured covariance matrices.

DISCUSSIONS

Using sample data from a cohort of older medical patients with multi-morbidities and polypharmacy, we illustrated how the exposure-crossover design can be extended to longitudinal data with repeated measures of both exposure and outcome in the quasi-experimental context. The methodological renovations of this longitudinal exposure-crossover approach included: 1) defining exposure-crossover intervals under biologically plausible assumptions for exposure window (avoiding statistical artifacts, serendipity and spurious association); 2) shifting focus from *subsequent-baseline* comparison to *subsequent-induction* comparison (to address the "causal" effect of medication treatment on an acute disease); 3) isolating *treatment episodes* from both the *induction* and *subsequent* intervals as a nuisance parameter (mitigating causal ambiguity); and 4) addressing proximal confounding *via* multivariable adjustment while controlling stable baseline risk through self-matching. Through this longitudinal exposure-crossover exploration, we observed a potential modest benefit of haloperidol treatment on reducing delirium severity before adjusting for ICU confounding, which is consistent with some clinical trials that demonstrated a benefit of low dose haloperidol treatment (<4.5 mg per day) on reducing delirium severity among post-surgery patients, but not the delirium incidence [6].

However, to appropriately accomplish the above goals is not trivial and requires deep critical thinking of the underlying causal mechanism and labor-intensive data-collection effort. Although utilizing repeated measure data increases the statistical efficiency or estimation precision, when many patients have missing data due to mortality or dropout, comparisons made across person-time segments may compromise the self-matching capacity in control for stable baseline risk. Therefore, potential survivor bias is likely and should be addressed with complete follow-up data allowing equal length for the two comparison intervals [1–3]. In addition, it remains to be clarified whether the diminishing haloperidol effects in this study reflected the overwhelming confounding unique to ICU setting (e.g., mechanical ventilation), or the reduced power due to excluding treatment-irrelevant person-time data.

Unlike other studies in the field that often focused on disease incidence, we modeled delirium severity as a quantitative outcome. Other than finer granulation on the scale, we were concerned of potential "treatment"-selection bias [1,16,17], because all patients in this sample were diagnosed with delirium. Modeling the incidence or recurrence of a diseases that potentially indicated the treatment at the first place may be subject to increased risk of "reverse causality". Indeed, a recent report from this EPIC cohort observed a significant association between repeated haloperidol doses and the next-day delirium incidence using marginal structural model [7].

CONCLUSIONS

To conclude, we have demonstrated steps of implementing exposure-crossover design to longitudinal data with repeated measures of both exposure and outcome in a sample of older medical ICU patients. The preliminary evidence suggested that this longitudinal extension is feasible and reasonably attains to the essential feature of self-matching, as evidenced by the trivial change in the alleged haloperidol treatment effects after adjusting for stable baseline confounding alone, material change when ignoring self-matching in the analyses, and relative invariance when including the treatment episodes as an explicit parameter or restricting to the first induction and the last subsequent intervals only. Future longitudinal exposure-crossover studies should consider using larger sample sizes, allowing wash-out period between successive treatment episodes and addressing informative missing data due to mortality and dropouts using alternative statistical models [1–3, 9, 16, 17].

Acknowledgments

This work was supported in part by the American Lung Association and Connecticut Thoracic Society (ID# CG-002-N), Claude D. Pepper Older Americans Independence Center at Yale University School of Medicine (2P30AG021342-06), the T. Franklin Williams Geriatric Development Initiative through The CHEST Foundation, ASP, Hartford Foundation, and grants through the National Institute on Aging [K23AG23023 and 1R21NR011066 (MAP) and R01 AG047891, R21 AG045148 (HGA)]. Drs. Han and Allore are supported by the Claude D. Pepper Older Americans Independence Center at Yale University School of Medicine (#P30AG021342 NIH/NIA).

References

1. Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. Am J Epideiol. 1991; 133(2):144–53.
2. Rothman, KJ.; Greenland, S. Modern epidemiology. 2. Philadelphia: Lippincott Williams & Wilkins; 1998. p. 111-112.p. 137
3. Redelmeier DA. The exposure-crossover design is a new method for studying sustained changes in recurrent events. J Clin Epidemiol. 2013; 66:955–63. http://dx.doi.org/10.1016/j.jclinepi.2013.05.003. [PubMed: 23850556]
4. Redelmeier DA, Zung JD, Thiruchelvam D, Tibshirani RJ. Fibromyalgia and the risk of a subsequent motor vehicle crash. J Rheumatol. 2015; 42:7.10.3899/jrheum.141315
5. Schlenker MB, Thiruchelvam D, Redelmeier DA. Intravitreal anti–vascular endothelial growth factor treatment and the risk of thromboembolism. Am J Ophthalmol. 2015; 160(3):569–580. http://dx.doi.org/10.1016/j.ajo.2015.06.011. [PubMed: 26116264]
6. Lonergan, E.; Britton, AM.; Luxenberg, J. Antipsychotics for delirium; Cochrane Database of Systematic Reviews. 2007. p. Art. No.: CD005594http://dx.doi.org/10.1002/14651858.CD005594.pub2
7. Pisani MA, Bramley K, Vest MT, Akgün KM, Araujo AL, Murphy TE. Association of cumulative dose of haloperidol with next-day delirium in older medical ICU patients. Crit Care Med. 2015; 43(5):996–1002. http://dx.doi.org/10.1097/CCM.0000000000000863. [PubMed: 25746748]
8. Goodman and Gilman's the pharmacological basis of therapeutics. 8. New York: Pergamon Press; 1990. p. 3-32.
9. Rubins DB. For objective causal inference, design trumps analysis. Ann Appl Statistics. 2008; 2(3): 808–840. http://dx.doi.org/10.1214/08-AOAS187.
10. Pisani MA, Murphy TE, Van Ness PH, et al. Characteristics associated with delirium in older patients in a medical intensive care unit. Arch Intern Med. 2007; 167:1629–1634. http://dx.doi.org/10.1001/archinte.167.15.1629. [PubMed: 17698685]

11. Ely EW, Margolin R, Francis J, et al. Evaluation of delirium in critically ill patients: validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). Crit Care Med. 2001; 29(7):1370–1379. http://dx.doi.org/10.1097/00003246-200107000-00012. [PubMed: 11445689]
12. Inouye SK, Kosar CM, Tommet D, et al. The CAM-S: Development and validation of a new scoring system for delirium severity in 2 cohorts. Ann Intern Med. 2014; 160:526–533. http://dx.doi.org/10.7326/M13-1927. [PubMed: 24733193]
13. Han L, McCusker J, Cole M, Abrahamowicz M, Primeau F, Elie M. Use of medications with anticholinergic effect predicts clinical severity of delirium symptoms in older medical inpatients. Arch Intern Med. 2001; 161:1099–105. http://dx.doi.org/10.1001/archinte.161.8.1099. [PubMed: 11322844]
14. Durán CE, Azermai M, Vander Stichele RH. Systematic review of anticholinergic risk scales in older adults. Eur J Clin Pharmacol. 2013; 69(7):1485–96. http://dx.doi.org/10.1007/s00228-013-1499-3. [PubMed: 23529548]
15. McCulloch, CE.; Searle, SR. Generalized, linear and mixed models. New York: John Wiley & Sons; 2001.
16. Maclure M, Mittleman MA. Should we use a case-crossover design? Annu Rev Public Health. 2000; 21:193–221. http://dx.doi.org/10.1146/annurev.publhealth.21.1.193. [PubMed: 10884952]
17. Stukel TA, Fisher ES, Wennberg DE, Alter DA, Gottlieb DJ, Vermeulen MJ. Analysis of observational studies in the presence of treatment selection bias. JAMA. 2007; 297:278–285. http://dx.doi.org/10.1001/jama.297.3.278. [PubMed: 17227979]

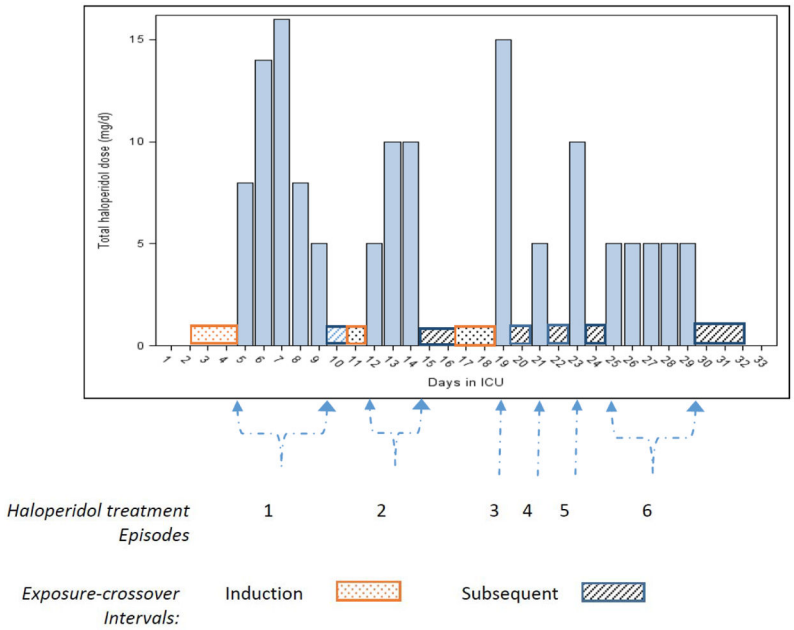


Figure 1.

A case scenario illustrating exposure-crossover sampling from a patient with 33 person-days of ICU stay.

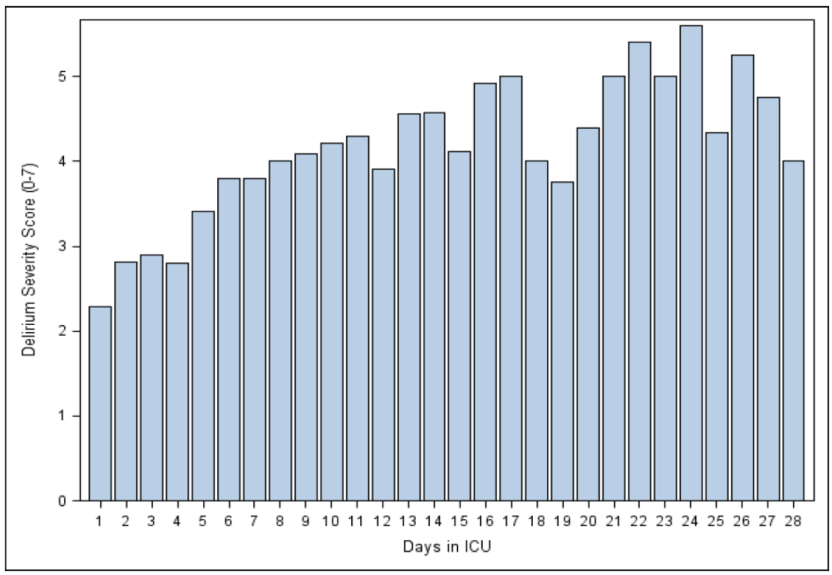


Figure 2.

Delirium severity score as measured with the CAM-ICU during ICU stay truncated at 28 days after admission.

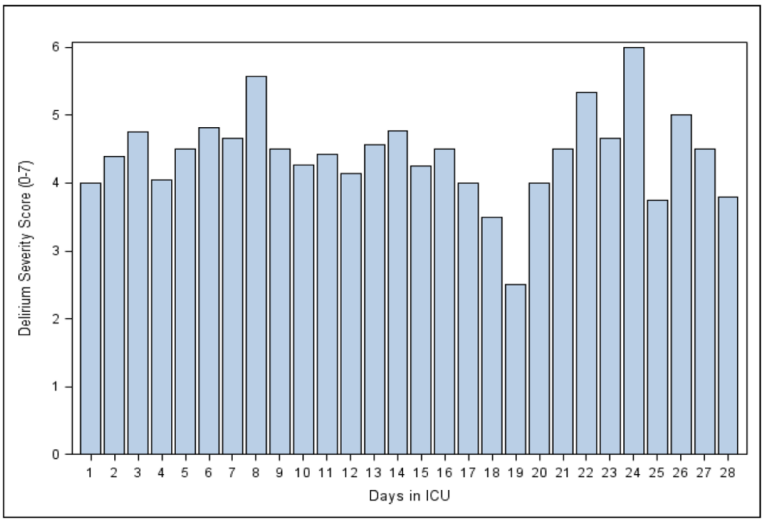
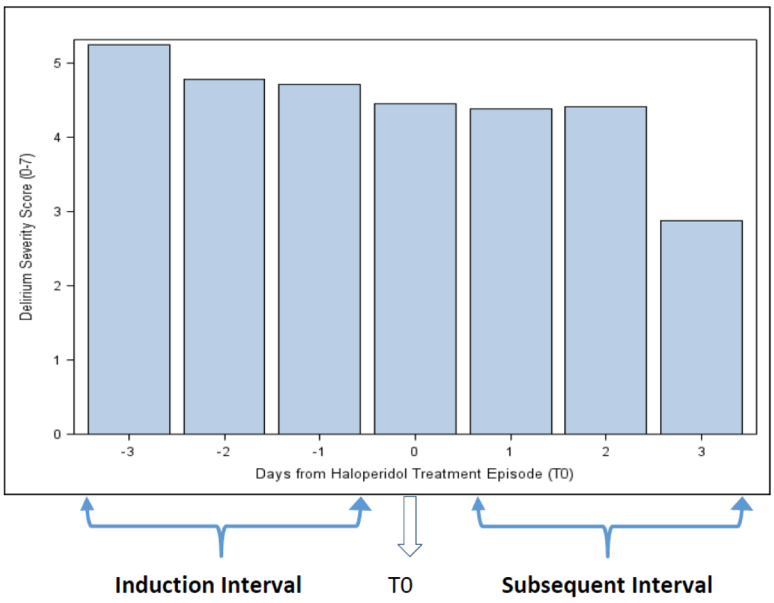


Figure 3.

Exposure-crossover sample: Delirium severity score as measured with the CAM-ICU during ICU stay truncated at 28 days after admission.



# person-days	48	64	88	-	130	60	37
Total		200		(398)		227	
#CAM-ICU scores	8	15	22	-	53	26	11
Total		45		(116)		90	

Figure 4.

Exposure-crossover sample: Delirium severity score as measured with the CAM-ICU aggregated over the *induction* and *subsequent* intervals across person-time.

Table 1

Characteristics of Study Sample

Characteristics	Full Cohort (N=304)	Analytic Sample (N=92)
Baseline characteristics		
Age (yr), mean±sd	74.7±8.5	74.3±7.7
Male gender, n (%)	143 (47.0)	41 (44.6)
IQCODE>3.3, * n (%)	94 (31.2)	37 (40.2)
APACHE II Score, † mean±sd	23.4±6.4	24.5±6.1
On antipsychotics, n (%)	20 (6.6)	7 (7.6)
Admitting diagnosis, n (%)		
Respiratory diseases	153 (50.3)	54 (58.7)
Gastrointestinal hemorrhage	52 (17.1)	8 (8.7)
Sepsis	50 (16.5)	14 (15.2)
Other diagnoses ‡	45 (16.1)	16 (17.4)
ICU characteristics		
Time under mechanical ventilation, n (%)		
Never intubated	144 (47.4)	39 (42.4)
Intubated <80% time	85 (28.0)	31 (33.7)
Intubated ≥80% time	75 (24.7)	22 (23.9)
Cumulative anticholinergic drug burden per day during ICU, § mean±sd	1.1±2.0	1.3±2.1

Abbreviations: sd, standard deviation; CAM, The Confusion Assessment Method; IQCODE, Informant Questionnaire on Cognitive Decline in the Elderly; APACHE, Acute Physiology and Chronic Health Evaluation; ADL, activities of daily living.

* Indicative of dementia.

† Acute Physiology and Chronic Health Evaluation II Score,

‡ Include neurological diseases, diabetes, metabolic abnormalities, acute renal failure and cardiac causes.

§ Represents cumulative dose standardized on WHO Defined Daily Dose for adults across 12 common anticholinergic medications received each day, including amitriptyline, atropine, dicyclomine, diphenhydramine, imipramine, benzodiazepine (lorazepam), meclizine, olanzapine, paroxetine, promethazine, and narcotics (fentanyl and morphine).

Table 2

Predicted Mean Difference of Delirium Severity Scores between the *subsequent* and *induction* Intervals in Exposure-Crossover Sample (N=92)

Model no./Covariates*		Mean Difference§ (95% CI)	Δ (%)¶	Comments
1	Unadjusted	−0.55 (−1.10, −0.01)	(reference)	Theoretically free of stable, baseline confounding due to self- matching
2	Adjusted for baseline confounding only†	−0.60 (−1.15, −0.04)	7.2	Theoretically redundant or over-adjusted after self-matching
3	Adjusted for ICU confounding only‡	−0.38 (−0.99, 0.24)	32.4	Theoretically unbiased from both baseline (*via* self-matching) and time-varying, proximal confounding (*via* regression adjustment)
4	Adjusted for both baseline and ICU confounding†‡	−0.38 (−1.01, 0.25)	31.9	Unbiased yet less efficient due to potential over-adjustment of baseline confounding

*Estimated using a generalized estimating equation model of the CAM-ICU-based delirium severity scores as a Poisson outcome, accounting for self-matching as a cluster and each persons' unique correlation pattern as unstructured.

†Included age in years, APACHE II score (continuous), admission diagnoses (4 categories), dementia (yes vs no) and antipsychotic use prior to ICU admission (yes vs no).

‡Included the time spent in mechanical ventilation during ICU stay (0, 1–80%, ≥80%) and a time-varying covariate, cumulative anticholinergic drug burden received each day.

§Represents predicted delirium severity score difference between the *subsequent* and *induction* Intervals using the GEE model.

¶Represents percent change in the effect estimates (Mean Difference) between each adjusted models versus the unadjusted (reference) model.

Table 3

Sensitivity Analyses: Adjusted Mean Difference of Delirium Severity Scores between the *Subsequent* and *Induction* Intervals Using Alternative Approach (N=92)

Model no./Alternative Approach*		Mean Difference† (95% CI)	Comments
1	Including haloperidol treatment episodes as an explicit parameter	−0.44 (−1.07, 0.19)	Avoid pragmatics and arbitrary exclusion of person-time data; may complicate interpretation due to potential reverse-causality
2	Assuming 1-day exposure window‡	−0.26 (−1.08, 0.55)	Address unequal lengths of comparison intervals (i.e., *subsequent* vs *induction*); may compromise model efficiency
3	Restricting to the last *subsequent* and first *induction* intervals from each patient	−0.51 (−1.24, 0.22)	Remove potential residual time-varying confounding and reverse-causality between successive exposure-crossover intervals at sacrifice of sample size
4	Without accounting for self-matching	−0.16 (−0.68, 0.37)	Reduce model complexity by ignoring treatment sequencing; may underestimate treatment effect

*All models were estimated using a generalized estimating equation Poisson model of the CAM-ICU-based delirium severity scores, accounting for self-matching and each persons' unique correlation pattern, except otherwise indicated.

†Represents predicted delirium severity score difference between the *subsequent* and *induction* Intervals from each GEE model, adjusted for 5 baseline (age, APACHE II score, admission diagnoses, dementia and antipsychotic use prior to ICU admission) and 2 ICU (percent time spent in mechanical ventilation during ICU stay and cumulative anticholinergic burden received each day) covariates.

‡Retained only the *induction* and *subsequent* intervals with 1-day duration in the analyses.