Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.JournalofSurgicalResearch.com

Letter to the editor

The use of coarsened exact matching to evaluate treatment mode in the rib fracture patient

We read with great interest the recent article by McKendy *et al.* titled "Epidural analgesia for traumatic rib fractures is associated with worse outcomes: a matched analysis."[1] In it, the authors found thoracic epidural analgesia (TEA) to elicit poorer outcomes than systemic analgesia among rib fracture patients. The keenness of our interest results from the rarity of this finding.

The underlying question—whether TEA is beneficial for rib fracture patients—continues to be explored by numerous authors to no satisfying end.[2] Many of the articles that fail to find support for TEA in this population are merely finding it to be equivalent to systemic treatments; seldom is TEA reported to be worse.[2] The lack of consensus in the effect of TEA is, in part, a consequence of discrepant study designs and statistical models employed to answer the same question.

In the study by McKendy *et al.*, the authors use coarsened exact matching (CEM) to create patient samples matched for demographic data and injury severity. They then conduct t-tests to detect significant differences in the samples created by this statistical tool. After matching for age, sex, five characterizations of injury severity, insertion of a chest tube, and year of injury, the authors found the TEA group to have 31% longer hospital stays ($P = 0.026$), a 90% greater incidence of pulmonary complications ($P = 0.009$), and nonsignificant patterns for TEA patients to fair worse in pneumonia (9% *versus* 5%; $P = 0.073$) and mortality (5% *versus* 2%; $P = 0.159$).

As noted, these findings piqued our interest, begging an investigation into the method used to obtain them. The CEM sampling procedure immediately stood out because it, too, appears to be rarely used. In our review of the literature, we were unable to find another study that employed CEM to evaluate different treatments among rib fracture patients. In fact, most of its applications seem to be in the social sciences, political science in particular.[3] So while we applaud the authors' sample design ingenuity, given their findings, we want to highlight some of the more subtle issues involved with the method.

CEM is most commonly employed to address the sampling problem of inferring a causal link between a treatment and the outcome of that treatment. The problem is that only one outcome is observed for each patient, who either did or did not receive the treatment, so one cannot observe a given patient's outcome if the treatment assignment had been different. This creates the possibility that treatment assignment is endogenous, and thus not random, which would bias the estimated treatment effect.

CEM attempts to mitigate this endogeneity by assuming a vector of observed covariates (X), which explains both the outcome and the treatment assignment. The authors use, among others, patient age for this purpose. They create different age strata, then "match" treated and controlled (nontreated) patients within each strata so that patient age is similar for both groups. Similar matching is done for the other covariates in X. Patients, both treated and not, who do not "match" are discarded. CEM thus culls the original sample to create what is referred to as balance in the new one. This smaller balanced sample then contains only treated and nontreated patients with values of X that are the same (or at least similar). Because X, therefore, does not affect treatment assignment, the estimation method used produces unbiased estimates of the outcome.

Although intuitively appealing, CEM has one clear shortcoming: it does not account for unobserved covariates that affect both the outcome and the treatment assignment.[3] Matching based solely on X, therefore, does not eliminate treatment endogeneity caused by something unobservable, and thus does not eliminate bias in estimated outcomes (e.g. something said in a conversation with a patient that may, consciously or subconsciously, affect a medical doctor's decision to treat).

As noted, CEM is a sampling tool; it is not a statistical estimation method. The authors' method compares means via t-scores. Because CEM has no obvious advantage over regression, we believe that binary logistic regression is a preferred method to predict mortality. The addition of the observed covariates to the regression does not allow them to bias the outcome or the treatment assignment. Moreover, it does so without having to cull the sample or having to define, somewhat arbitrarily, the strata used for the matching.

DOI of original article: https://doi.org/10.1016/j.jss.2017.02.057

Regarding this last point, the authors describe their definition of age strata and explain that CEM creates less bias than propensity score matching, yet they do not explore how sensitive their strata definitions may be to alternatives. We think this is worth mentioning because our results, estimated with and without CEM and using regression, described below, are quite different.

We are not suggesting that regression eliminates bias from treatment assignment endogeneity. The error term implicit in all regression modeling contains the same unobservable covariates that CEM ignores. We are simply saying that regression is preferred because it avoids the arbitrary definitions of the strata used for matching and it uses more information to predict outcomes because there is no culling of the sample. Of course, the surest way to eliminate estimation bias of the sort discussed here is to randomly assign patients to treatment and control groups before any estimation is undertaken. A large enough sample will then, on average, uniformly distribute both observed and unobserved covariates among both groups.

At our institution, we have a patient registry with a similar number of patients ($n = 1344$) treated over a 5-year period. We replicated the CEM analysis, which yielded a new sample of 722 matched subjects (144 patients received TEA and 578 did not). In this CEM sample, overall mortality was 4.3%, and all but one of these cases were in the non-TEA group; this difference was significant ($P = 0.017$). We also replicated the authors' subsample analyses and found TEA to be associated with lower rates of mortality among patients with ≥ 3 rib fractures ($P = 0.014$) and among patients with Injury Severity Scores ≥ 16 ($P = 0.022$).

After replicating the authors' study design, we also conducted logistic regression on our CEM sample, using the matching variables as predictors. In doing so, we found the odds ratio of TEA for mortality to be 0.053 ($P = 0.008$). Similar outcomes were found in the subgroup analyses.

We applaud the author's ingenuity in statistical design and encourage other authors to proceed with these ideas in mind. We also encourage other authors to employ regression analysis on the matched samples for proper control over confounders.

Acknowledgment

Authors' contributions: All authors participated in the conception of the project. The manuscript was written by M.M.A., W.E.H., and C.D.J. and was reviewed, edited, and given final approval by all authors. The patients were analyzed by M.M.A. and C.D.J. Oversight of data collection was done by J.M.S., L.E.H., and C.D.J. The statistical assessments were made by W.E.H., M.M.A., and C.D.J. All authors vouch for the accuracy and integrity of the data recorded, analyzed, and reported.

Disclosures

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

1. McKendy K, Lee L, Boulva K, et al. Epidural analgesia for traumatic rib fractures is associated with worse outcomes: a matched analysis. *J Surg Res*. 2017;214:117–123.
2. Galvagno S, Smith C, Varon A, et al. Pain management for blunt thoracic trauma: a joint practice management guideline from the Eastern Association for the Surgery of Trauma and Trauma Anesthesiology Society. *J Trauma Acute Care Surg*. 2016;81:936–951.
3. Miller M. *The uses and abuses of matching in political science*. Washington, DC: POLMETH, George Washington Univesity; 2013.

Michelle M. Amaral, PhD
William E. Herrin, PhD
Economics Department, University of the Pacific, Stockton, California

Jonathan M. Saxe, MD
Lewis E. Jacobson, MD
St. Vincent Hospital, Trauma Center, Indianapolis, Indiana

Courtney D. Jensen, PhD*
University of the Pacific, Stockton, California

**Corresponding author*. HES Department, University of the Pacific, 3601 Pacific Avenue, Stockton, CA 95211. Tel.: +1 209 946 3133; fax: + 1 209 946 3225.
E-mail address: cjensen1@pacific.edu

0022-4804/$ – see front matter

https://doi.org/10.1016/j.jss.2017.11.028