



# Addressing unmeasured confounding bias with a prior knowledge guided approach: coronary artery bypass grafting (CABG) versus percutaneous coronary intervention (PCI) in patients with stable ischemic heart disease

Lewei Duan<sup>1</sup> · Ming-Sum Lee<sup>2</sup> · Jason N. Doctor<sup>3</sup> · John L. Adams<sup>4</sup>

Received: 18 January 2022 / Revised: 6 June 2022 / Accepted: 13 June 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

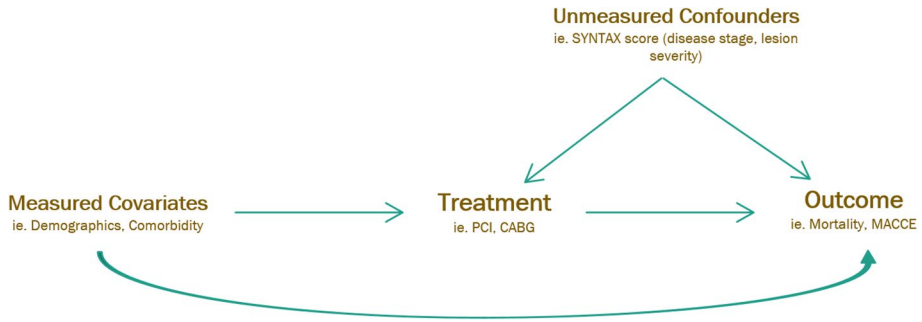
Unmeasured confounding undermines the validity of observational studies. Although randomized clinical trials (RCTs) are considered the “gold standard” of study types, we often observe divergent findings between RCTs and empirical settings. We present the “L-table”, a simulation-based, prior knowledge (e.g., RCTs) guided approach that estimates the true effect adjusting for the potential influence of unmeasured confounders when using observational data. Using electronic health record data from Kaiser Permanente Southern California, we compare the effectiveness of coronary artery bypass grafting (CABG) and percutaneous coronary intervention (PCI) on endpoints at 1, 3, 5, and 10 years for patients with stable ischemic heart disease. We applied the L-table approach to the propensity score adjusted cohort to derive the omitted-confounder-adjusted estimated effects. After the L-table adjustment, CABG patients are 57.6% less likely to encounter major adverse cardiac and cerebrovascular event (MACCE) at 1 year (OR [95% CI] 0.424 [0.396, 0.517]), 56.4% less likely at 3 years (OR [95% CI] 0.436 [0.369, 0.527]), and 48.9% less likely at 5 years (OR [95% CI] 0.511 [0.451, 0.538]). CABG patients are also 49.5% less likely to die by the end of 10 years than PCI patients (OR [95% CI] 0.505 [0.446, 0.582]). We found the estimated true effects all shifted towards CABG as a more effective procedure that led to better health outcomes compared to PCI. Unlike existing sensitivity tools, the L-table approach explicitly lays out probable values and can therefore better support clinical decision-making. We recommend using L-table as a supplement to available techniques of sensitivity analysis.

**Keywords** Unmeasured confounding · Bias · Simulation-based · Prior knowledge guided approach · Percutaneous coronary intervention (PCI) · Coronary artery bypass grafting (CABG)

---

✉ Lewei Duan  
lewei.duan@kp.org

Extended author information available on the last page of the article



**Fig. 1** Causal relationship between treatment, outcome, and confounders

## 1 Introduction

Healthcare utilization databases represent a major data source for studies seeking to infer causal relationships between treatments and outcomes. The principal challenge for such studies comes from unmeasured confounders (Fig. 1). It is important that observational studies address the issue of unmeasured confounding as it is an often-overlooked threat to validity of observational studies (Greenland 1996, 2009, 1999, 2005; Imbens and Rubin 2015; Rosenbaum 2010; VanderWeele and Ding 2017). With observational data, unmeasured confounding bias is a central limitation—uncontrolled and unmeasured covariates may confound the relationship between treatment and outcome (Rosenbaum 1991; Cornfield et al. 1959).

A comprehensive array of sensitivity analysis techniques has been designed to evaluate evidence of causation in the presence of unmeasured confounding (Greenland 1996, 1999; VanderWeele and Ding 2017; Rosenbaum and Small 2017; Greenland and Mansournia 2015; Chiba 2012; Vanderweele and Arah 2011; Schneeweiss 2006; Brumback et al. 2004; Rosenbaum and Rubin 1983). Sensitivity analysis typically assess the magnitude of possible biases, and report the level of confidence for the study results.

As a supplement to existing sensitivity analysis, the E-value was introduced to assess the strength of the unmeasured confounding that could explain away the effects, thereby casting reasonable doubt on the accuracy of the estimate (VanderWeele and Ding 2017). E-value was defined as “the minimum strength of association, on the risk ratio scale, that an unmeasured confounder would need to have with both the treatment and the outcome to fully explain away a specific treatment–outcome association, conditional on the measured covariates” (VanderWeele and Ding 2017). In general, a small magnitude of E-value should increase concern that unmeasured confounding bias might be substantial relative to the estimated effect. Since its introduction, E-values have helped alleviate the over-reliance on the p-value and offset inadequate assessments of robustness to bias. Nevertheless, the E-value derivation relies on assumed relationships between binary variables, which do not necessarily generalize to mixed binary and continuous variables. The discussion around E-values inspired us to develop a new tool to supplement previously uncovered areas in

sensitivity analysis of unmeasured confounders. The method we propose aims to help researchers find a range of likely true values rather than bounding the true value or focusing on statistical significance. Presenting a range rather than simply a bounding end point gives a more balanced sense of likely values.

We use an empirical example to illustrate how this new tool works. The choice of optimal coronary revascularization methods has been vigorously debated: CABG is more invasive than PCI, has a longer recovery time, remains more costly, but the incremental cost-effectiveness ratio was favorable (Lim 2014). The SYNTAX score is a scoring system calculated based on clinical elements indicating disease stages. While it is a useful proxy for disease severity, the severity of coronary artery disease is often evaluated by physicians' visual inspection, and the SYNTAX score is rarely calculated and thus remains unmeasured in electronic health records (EHR). Several randomized clinical trials (RCTs) found that patients with low and intermediate SYNTAX scores can be treated with PCI or CABG with equal results; and those with a high score do better with CABG (Mohr et al. 2013; Serruys et al. 2009; Kappetein et al. 2011; Thuijs et al. 2019). However, to our knowledge no observational studies have attained sufficient sample size and hence power, to detect the difference.

RCTs were often considered the “gold standard” of study types. Nevertheless, we often observe divergent findings between RCTs and empirical settings (Rothman 2014; Hernan et al. 2013). Rather than simply summarize the literature from RCTs as an evidence report, we consider how the answers might be different in our situation, and whether the conclusions drawn from RCTs can be generalized to our circumstances. In this paper we show how we use information from both settings to reach a conclusion. We make no claim that the estimate of treatment effect we generate is better than an RCT for the general population. We simply suggest that the proper estimate of the treatment effect for our population might be different from what one would see in a clinical trial. Consequently, we seek to incorporate the internal validity of a clinical trial that may not directly generalize to our unique system and population.

In this study we compare the safety and effectiveness of CABG vs. PCI for patients with stable ischemic heart disease using observational data, and introduce the “L-table,” a simulation-based, prior knowledge guided approach, which enables investigators to use external information about unmeasured confounders (e.g., SYNTAX score) to identify a plausible range of estimated true effects in observational studies. We call it the L-table, named after uncovering the “Location” of the graphical frame in the numeric table that contains the estimated true effects corresponding to assumed correlations. Our framework is flexible enough to adapt to different types of models. The statistical code for L-table is open source with customizable parameters.

This approach uses well-known relationships from the literature. Although much of the literature is focused on testing statistical significance, our focus falls on understanding plausible effect sizes. This is an often-underappreciated element in applying confounder analysis to practical decision-making. Some consideration of the practical significance of the results must follow statistical significance. This is especially true in comparative effectiveness problems where an incumbent treatment already exists. A focus on plausible effect sizes better supports some real-world applications like cost effectiveness analysis and the allocation of scarce resources to competing interventions.

## 2 Methods

### 2.1 Model framework

We assume that the effect of a treatment on the outcomes can be estimated by a generalized linear model. In this study, we use a logistic model to illustrate how we have used the proposed method to evaluate the effectiveness of cardiovascular procedures. The dependent variable is a binary outcome of interest; the independent variables include treatment, observed and unobserved confounders.

It is important to be precise in the way we talk about correlations when applying this method. The main issue is the difference between an *estimator* and an *estimand*. An *estimator* uses a computation on observed data to calculate an estimate. Here we will use *estimator* for parameters, typically unknown, that we wish to estimate. An easy source of confusion in our setting is that an estimator may be estimating different things depending on the type of data it is applied to. For example, the Pearson correlation (*estimator*) estimates the product-moment correlation (*estimand*) when applied to bivariate normal data. The same estimator calculates the Spearman correlation when applied to binary data.

Our simulation framing uses a trivariate normal distribution to generate the data that we summarize to estimate the bias from unmeasured confounders. The correlations are used to construct a simulation data set with a known true value. Correlation-based simulation is the simplest way to characterize bias in a normal theory model. Specifically, the correlations between unmeasured confounder and other components in the model were specified with assigned values to simulate the dataset generating the unmeasured confounding bias. This is one standard approach to characterizing omitted variable bias (Maddala 1983). It is a more flexible and general approach than the manipulation of odds or risk ratios from contingency tables. We will refer to these correlations as “product-moment” correlations (or “latent” correlations). This is similar to the way many econometrics books motivate limited dependent variables (Maddala 1983). The challenge is that correlations available in the literature or that inform clinical intuition may not be straight-forward estimates of the product-moment correlations. It may be necessary to transform the available correlations into these product-moment correlations to run the simulations. After the simulations have been run it may be necessary to transform the results back to the original scale for reporting in the original context.

#### (1) The tetrachoric correlation case

Tetrachoric correlation characterizes the data in a 2 by 2 table by hypothesizing a latent bivariate normal variable that is dichotomized in both dimensions to generate the 2 by 2 table. The simple Spearman correlation calculated from the 2 by 2 table is not an estimate of the latent correlation. What is needed is a tetrachoric correlation estimate.

This will be the common case for the relationship between treatment and outcome. Fortunately, this is data we will have in observational studies. Getting the needed latent correlation estimate is simply a matter of calculating the tetrachoric correlation with appropriate software. If the original Spearman metric is desired for reporting results the dichotomization that defines the tetrachoric correlation can be mimicked in the simulation.

The generalization of tetrachoric correlation to ordered categorical variables in one or both dimensions is called polychoric correlation. If ordered categorical omitted

confounders are available in the literature the same process is used with polychoric correlations instead of tetrachoric correlations.

(2) The point-biserial case

Point-biserial correlation is the application of the Pearson correlation calculation to a binary and a continuous variable. Conceptually, our framework is similar to the tetrachoric correlation situation, but in this case only one of the two latent variables has been dichotomized. We want to use the latent correlation in our simulation, but the point-biserial correlation does not estimate the latent correlation. However, the biserial correlation (Cox 1974) provides an estimate of the latent correlation we need.

This is the case we commonly need to address when we estimate the correlation between the treatment and a continuous unmeasured confounder. This estimate usually comes from the literature where we are unlikely to have access to the subject level data. Fortunately, there is a formula to transform the point-biserial correlation into a biserial correlation using only summary quantities that are often available in the literature. As in the tetrachoric correlation case, dichotomization in the simulation can transform the results to the original point-biserial scale if needed as an aid to interpretation.

Our method simulates from these underlying correlations the drivers of the causal relationship to assess the estimated true effect when an important confounder is omitted. The model framework can be illustrated in a pipeline composed of four major steps (Fig. 2).

Step 1. Model empirical data.

1. (Optional but Optimal) Balance the empirical data. We used matched propensity score weighting (other propensity score methods can be substituted).
2. Estimate the effect based on balanced empirical data:  $\beta_{est\_emp}$
3. Calculate proportion of treatment and outcome: *treatment (%)* and *outcome (%)*

Step 2. Ascertain correlation coefficients.

1. Calculate correlation between outcome and treatment ( $\rho_{YT\_emp}$ ) based on empirical data. In this study, we applied tetrachoric correlation for the binary outcome and binary treatment. In a log-normal cost model a biserial correlation could be used.

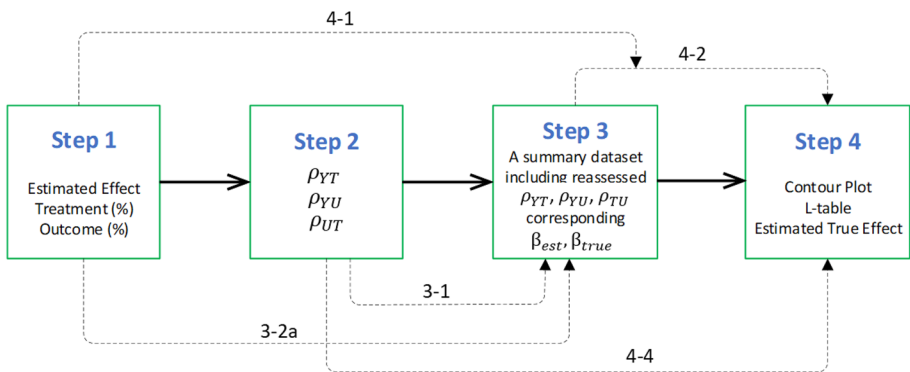


Fig. 2 Pipeline of model framework to generate L-table

2. Deduce observed correlations between the unmeasured confounder and outcome ( $\rho_{YU\_ext}$ ), and correlations between the unmeasured confounder and treatment ( $\rho_{TU\_ext}$ ) based on external information. We calculated both polychoric correlation (for simulation input) and biserial correlation (for matching L-table's label), because the SYNTAX score was reported in ordinal form in the RCTs, and its original form is continuous.
3. Make a sequence of values for each latent correlation coefficient estimate with small intervals. We created a sequence using  $\rho \pm 0.3$  with 0.05 intervals.

### Step 3. Simulation.

1. Construct a collection of correlation matrices, tables that contain correlation coefficients between variables. These matrices are restricted to the positive definite cases; and are based on all combinations of the latent correlation coefficients estimations (from Step 2–3).
2. Simulate multivariate normal distributed datasets based on the collection of correlation matrices. Each dataset is generated based on one correlation coefficient matrix.
  - (a) Simulate 100 (an adjustable parameter) datasets according to each correlation matrix. Each dataset contains 1000 (an adjustable parameter) data-points.
  - (b) Dichotomize the outcomes and treatments based on the proportions of treatment and outcome in empirical data (from Step 1–3).
  - (c) Reassess the correlation coefficients in simulated datasets based on the recently dichotomized variables, and result in biserial  $\rho_{YU\_sim}$  and biserial  $\rho_{TU\_sim}$  (estimates of the latent correlations).
  - (d) Perform suitable modeling for simulated datasets. We used logistic regression in this example.
  - (e) Output means of your estimates ( $\beta_{est\_sim}$ ) based on the model of each 100 datasets without inclusion of the measured cofounder, and the estimated true estimates ( $\beta_{true\_sim}$ ) with inclusion of the unmeasured confounder in the model (from Step 3-2d), along with the corresponding estimates of the latent correlation in simulated datasets (from Step 3-2c).

### Step 4. Obtain estimated true effects.

1. Select a small range around your empirical estimate  $\beta_{est\_emp}$  (from Step 1–2). We used  $\ln \beta_{est\_emp} \pm 0.1$ .
2. Subset the outputs (from Step 3-2e) based on the collections of values around  $\beta_{est\_emp}$  (from Step 4 -1) that match  $\beta_{est\_sim}$  (from Step 3-2e). This selected data includes  $\beta_{est\_sim}$ ,  $\beta_{true\_sim}$ ,  $\rho_{YU\_sim}$ , and  $\rho_{TU\_sim}$ .
3. Construct L-table
  - (a) (optional) Output contour plot
  - (b) Use the subgroup data (from Step 4–2) to tabulate  $\beta_{true\_sim}$  to form the L-table, which employs  $\rho_{TU\_sim}$  as column label and  $\rho_{YU\_sim}$  as row label.
4. Assess the estimated true effects

Based on L-table (from step 4-3b), we can locate the estimated true effect  $\beta_{true}$  using  $\rho_{YU\_ext}$  and  $\rho_{TU\_ext}$  in original context (from Step 2–2) to match the L-table's labels. We used biserial correlations in this study.

Confidence intervals of the estimated true effect can be obtained by repeating step 4 with confidence limits from the empirical analysis replacing the effect estimate.

## 2.2 Simulation calculation

The simulation uses a generalized linear model relationship between outcome and treatment that is adjusted for an unmeasured confounder. The latent outcome, treatment and unmeasured variables are generated by a multivariate normal distribution. The covariance matrix is created based on the correlation matrix which is composed of the correlation coefficients guided by prior knowledge (Supplementary Figure 1). Because the latent correlation will change to an observed correlation after some variables were dichotomized in simulation, we recommend users assign an adequate range surrounding the latent correlation coefficients to span the needed range of observed correlations after dichotomization. Within this range, we specify the values of correlation coefficients with equally spaced values.

The simulations only use correlation matrices that satisfied positive definiteness. By working with the correlation matrix, we are implicitly using standardized regression coefficients. In applications with continuous outcomes transformation to the original coefficient scale will be required. The propensity score adjustment with matching weights assures covariate balance between the treatment and comparison groups, and allows us to work with our simple trivariate normal setup in simulation—so that we do not need to simulate other confounders. We also assume that the unmeasured confounder used in the simulation represents the composite effect from all unmeasured confounders.

Once the initial dataset was simulated, we dichotomized outcome and treatment variables based on their corresponding proportions as found in the empirical data. We then calculated the latent correlation coefficients of outcome and unmeasured variable, and latent correlation coefficients of treatment and unmeasured variable, in order to use them in the results graph. These updated correlation coefficients were used to label the coordinates of the contour plot, or columns and rows for the L-table. We simulated each dataset with 1000 observations based on a corresponding correlation matrix. We made 100 iterations of this data generation procedure. Mean and median of the estimated effects with and without adjusting for unmeasured confounder were obtained. We used the multivariate normal approach to generate the data that can be easily generalized to other model types upon modification of the R code (Duan 2021).

## 2.3 Data source

Our analysis is based on a retrospective cohort study using EHR data from the Kaiser Permanente Southern California (KPSC) Health System. KPSC provides care to a racially, ethnically and socio-economically diverse population that is broadly representative of the racial-ethnic groups of Southern California (Derose et al. 2013). The study protocol was approved by the KPSC Institutional Review Board (IRB). A waiver of informed consent was obtained due to the observational nature of the study.

Using procedure codes and International Classification of Diseases (ICD) 9/10 codes (Supplementary Table 1), we searched EHR data for adult patients (age  $\geq 18$  years) who

underwent a revascularization procedure (CABG or PCI) between January 1st, 2006 and March 1st, 2015. We chose March 1st, 2015 as the end of the inclusion period to mitigate any impacts on hospitalization, mortality, and the means to facilitate revascularization procedures caused by the Covid-19 pandemic and to ensure this cohort has at least 5 years follow up time until March 2020, when healthcare utilization in California was affected by the state's stay-at-home order. The index date was defined as that of the first revascularization procedure the patient received. We excluded patients who were not KPSC members or who did not hold continuous one-year membership prior to the index date (allowing for a 45-days gap), and who underwent revascularization procedures prior to the index date. To identify patients with ischemic heart disease, two conditions were used: first, a principal diagnosis of Coronary Artery disease or Angina; second, no prior history of acute myocardial infarction. For inclusion in this cohort, both conditions had to be met at least twice in outpatient visits or at least once in inpatient admission within a year before the index date (Fig. 3).

Using this cohort, we created sub-cohorts for the 1-year, 3-year, 5-year, and 10-year endpoints to study the treatment effects in the short- and long-term. To ensure sufficient follow-up time for each patient, patients were excluded from the sub-cohort analysis if their KPSC membership lapsed prior to the end of each study period. Patients with insufficient follow-up time but experienced the endpoints wouldn't be captured by our data, hence they were excluded from the particular sub-cohort analysis. Patients who experienced an outcome event in early sub-cohorts will be considered as having an outcome event in the later sub-cohorts regardless of the change in their membership status. We moved forward the end of inclusion date for the 10-year cohort to March 1st, 2010 to allow sufficient follow-up time. Patients' disease history and outcomes were identified using the ICD 9/10 codes (Supplementary Table 2).

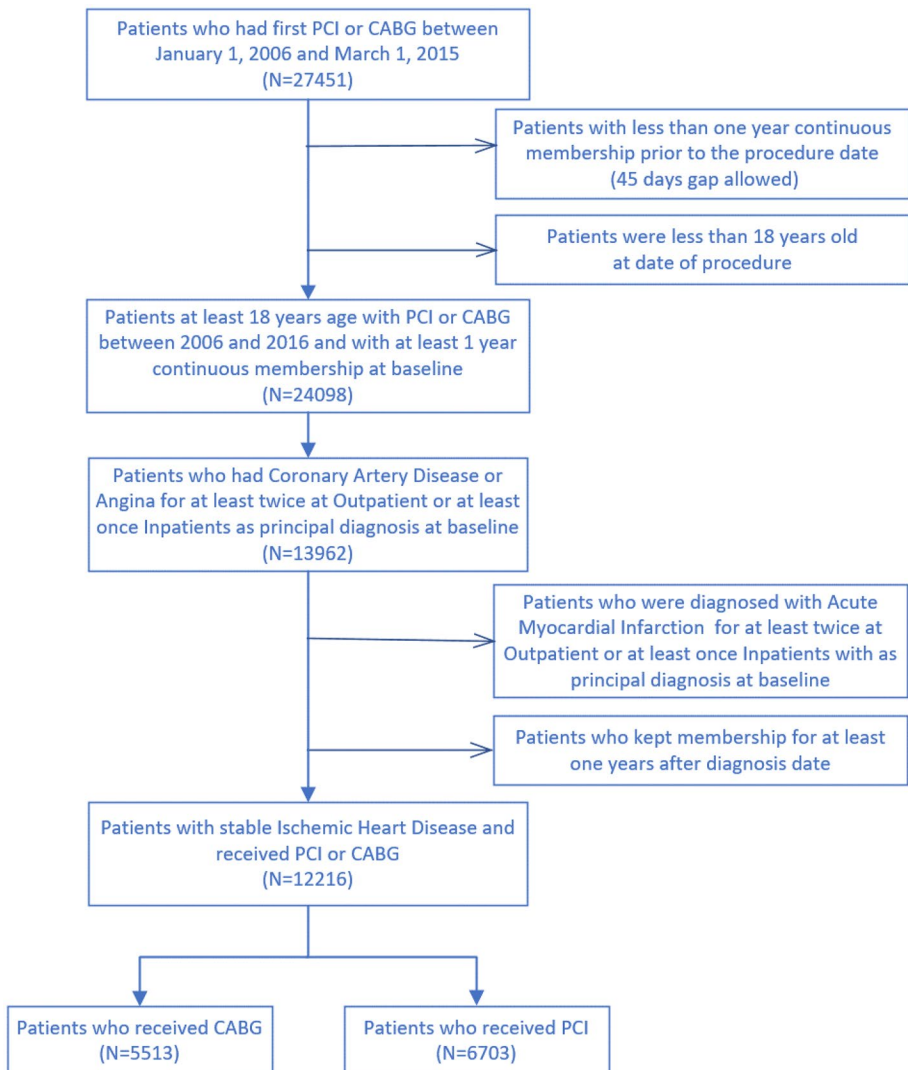
We identified covariates in the following categories: baseline demographics, medical comorbidities, cardiac risk factors, and cardiac medication usage. Medical comorbidities and cardiac risk factors were collected using ICD codes for one year prior to the index date. Baseline concomitant medications were identified using outpatient pharmacy records. The study endpoints include all-cause mortality, hospitalization from either myocardial infarction or stroke, repeat revascularization, and composite major cardiovascular events (MACCE) (Mohr et al. 2013), defined as one of above endpoints by the end of 1 year, 3 years, 5 years, or 10 years from the index date.

Mortality data was pulled from a mortality data mart derived from multiple sources: state of California's death master files, Social Security Administrative death master files, hospital death records, and insurance enrollment records. The endpoints for myocardial infarction and stroke were identified with a principal diagnosis at inpatient settings. The repeated revascularization procedures were identified as the event after discharge following the first procedure.

## 2.4 External information

To locate the estimated true effects using the L-table, we need to identify potential unmeasured confounders and to extract external information from available sources. The essential information concerns the correlation between unmeasured confounding and outcomes, and the correlation between unmeasured confounding and treatment. We need one L-table for each outcome. These correlation coefficients can be obtained from a pilot study or other relevant publications. Users may need to perform transformations or simulations to convert





**Fig. 3** Study cohort flowchart

the source information to suitable forms. For instance, based on available information, user can transform the point-biserial correlation into a biserial correlation; dichotomize continuous values in simulation to form a point-biserial scale.

We derived the correlation coefficients between a SYNTAX score and outcomes from published results of the SYNTAX trials (Mohr et al. 2013; Serruys et al. 2009; Kappetein et al. 2011; Thuijs et al. 2019). In this empirical example, we considered a SYNTAX score representing a linear combination of multiple omitted confounders. We chose to gather this information from RCTs because most sources of measured and unmeasured confounding are mitigated by the study design. In applications where a high quality RCT is not available, a plausible range of correlations may be estimated using clinical judgement. We derived the correlation between the SYNTAX scores and treatment from an observational

study (Valle et al. 2019) which accounts for the behavior effects of physicians' and patients' treatment choices. Subjective choices of this correlation may be used for sensitivity analysis or when suitable observational analyses are not available.

## 2.5 Statistical analysis

Descriptive statistics on patients' characteristics, including demographics, comorbidity, and medication history were reported by treatment groups using frequencies and percentages. We calculated standardized mean difference before and after applying matching weights. A difference of 0.10 or less was considered an adequate balance between the two groups.

To resemble the enrollment patterns in RCTs and adjust for selection bias, we prepared data with matching weight (Li and Greene 2013), a propensity score weighting method. Matching weights are a variant of the inverse probability weight, but the matching weight estimator assigns greater weight to individuals whose propensity score is close to 0.5 (i.e., the circumstance generated in a two-arm RCT with clinical equipoise). The underlying propensity score model used a logistic regression between treatment (dependent variable) and 57 baseline covariates (independent variables), which included age, sex, race, comorbidities and baseline medication use.

Logistic regression was performed to assess the treatment effects. We reported crude (unadjusted), and matching weighted Odds Ratios, along with their 95% confidence interval,  $p$  values, and E-values. A small E value suggested that unmeasured confounding should be a concern (VanderWeele and Ding 2017). A  $p$  value of  $<0.05$  was the nominal level of significance.

Using simulated datasets, we assessed estimated effects with and without the unmeasured confounder in the model. Using the method of bivariate interpolation for irregularly distributed data points (Akima 1978, 1996), we are able to plot the estimated true effects against biserial correlations coefficients between outcome and unmeasured confounding, and between treatment and unmeasured confounding. The pattern of the estimated true effect in relation to these correlations can be visualized in the contour plot. We tabulated these estimated true effects to form the L-table, similar to the assembly of a Chi-squared table, where the column was labeled with the correlation coefficients between treatment and unmeasured confounders, and the row was labeled with the correlation coefficients between outcome and unmeasured confounders. The user can adopt the L-table as a reference to locate a plausible range of estimated true effects with prior information about the correlations ( $\rho_{YU_{ext}}$  and  $\rho_{TU_{ext}}$ ). The values bounded by the perimeter within the L-table represent the estimated true values of the causal effect on the scale of the applied analytic model. This effect is aligned with the correlations between unmeasured confounding and outcome on the one hand, and between unmeasured confounding and treatment—conditional on the measured covariates—on the other.

Published studies may report varied findings. The polychoric correlation between treatment selection and the SYNTAX score derived from external publication we used was 0.625. To explore how sensitive the L-Table is to variation in inputs, we performed a sensitivity analysis by assigning two different correlation coefficients (0.5 and 0.7) as the relationships between the treatment and the unmeasured confounders in simulation, and ran the models with the same procedures. All statistical analyses were conducted in SAS 9.4 (SAS Institute Inc., Cary, NC) and R version 4.1.0 (R core Team 2021).

### 3 Results

We identified 12,216 adult KPSC members with stable ischemic heart disease who underwent PCI or CABG between January 1, 2006 and March 1, 2015. Among these patients, 5513 received CABG, and 6703 received PCI. In this group, 11,298 patients maintained their memberships by the 5th year: 5158 received CABG and 6140 received PCI (Table 1). The average age of CABG patients was 2.5 years older than their PCI counterparts ( $p < 0.001$ ). More men than women underwent CABG (77.5% vs. 73.5%,  $p < 0.001$ ). Asians were more likely to undergo CABG (11.6%) compared to PCI (9.0%). In general, CABG patients had more comorbidities and more frequently used cardiac medications. The baseline characteristics of the study cohort between CABG and PCI were balanced (SMD  $< 0.1$ ) after propensity-score adjustment with matching weights. We used the weighted cohorts for our analysis.

After propensity score adjustment, patients treated with CABG were less likely to experience MACCE, mortality, hospitalization for MI, or repeat revascularization (Table 2). However, CABG patients are more likely to be hospitalized due to stroke after the revascularization procedure, but these effects were not statistically significant. All E-values are small, which suggests that unmeasured confounding should be considered.

We assessed the proportion of outcome and the proportion of treatment from adjusted empirical data, and correlation coefficients between each of the end points and SYNTAX score (low, intermediate, high) based on the SYNTAX trials (Mohr et al. 2013; Serruys et al. 2009; Kappetein et al. 2011; Thuijs et al. 2019) (Supplementary Tables 3, 4). Using 5-year MACCE as an example, we first calculated the tetrachoric correlation coefficients between treatment and outcome from our data ( $\rho_{YT} = -0.047$ ); and derived the polychoric correlation coefficients based on an RCT study (Mohr et al. 2013) ( $\rho_{YU} = 0.062$ ) and a population-based study (Valle et al. 2019) ( $\rho_{TU} = 0.625$ ). We then assigned these parameters to generate simulated data, where the treatment (45.7%) and outcome (25.1%) variables would be dichotomized based on empirical data.

The contour plot (Fig. 4) shows the estimated true effects when the estimated effects from simulation were within a close range of the estimated effect for MACCE at 5 years (OR = 0.759). A point or range of true effects can be identified based on the correlation coefficients on the horizontal and vertical labels. This set of the estimated true effects were tabulated to form the L-table (Supplement Figure 2). We presented an example of a partial L-table (Table 3) to illustrate how we identify the estimated true effects based on the present latent correlation (biserial correlation):  $\rho_{TU} = 0.586$  and  $\rho_{YU} = 0.055$ . We took the mean of the values bordered by the frame and reached the desired Odds Ratio: 0.511.

After the adjustments made with the L-table, we found the estimated true effects all shifted in the direction that favors CABG as a much more effective procedure that led to better health outcomes compared to PCI, including hospitalization due to stroke at 5 year (OR [95% CI] 0.824 [0.649, 1.035]) (Table 2). These results suggest that the unmeasured confounder –disease severity, may have undermined the true contribution of CABG, because sicker people are more likely to receive CABG.

Different correlation inputs resulted in different L-table adjusted true effects. For 5-year MACCE, our original estimated true odds ratio was 0.511 (95% CI [0.451, 0.538]). When we assign  $\rho_{TU} = 0.5$  or 0.7 in simulation, the L-table adjusted true odds ratio became 0.45 (95% CI [0.397, 0.516]) and 0.517 (95% CI [0.459, 0.607]). These directions of the adjusted effect were the same. However, we found that hospitalization

**Table 1** Characteristics of study cohort included at 5 years endpoints

	Overall (N = 11,298)	CABG (N = 5158)	PCI (N = 6140)	SMD Unadjusted	SMD MV-weighted
<i>Demographics</i>					
Age (mean (SD))	67.48 (10.70)	68.84 (9.76)	66.33 (11.31)	0.238	0.002
GENDER: Male (%)	8513 (75.3)	4000 (77.5)	4513 (73.5)	0.094	0.003
Race/ethnicity (%)				0.118	0.006
Asian	1151 (10.2)	596 (11.6)	555 (9.0)		
Black	795 (7.0)	305 (5.9)	490 (8.0)		
Hispanic	2346 (20.8)	1064 (20.6)	1282 (20.9)		
Other	351 (3.1)	141 (2.7)	210 (3.4)		
White	6655 (58.9)	3052 (59.2)	3603 (58.7)		
Education: more than 75% high school graduate (%)	8527 (75.5)	3976 (77.1)	4551 (74.1)	0.069	0.001
Income: median household more than \$45,000 (%)	8301 (73.5)	3848 (74.6)	4453 (72.5)	0.047	0.006
Insurance type: medicare or medicaid (%)	3457 (30.6)	1724 (33.4)	1733 (28.2)	0.113	0.004
BMI (mean (SD))	29.03 (5.91)	28.76 (5.66)	29.25 (6.11)	0.083	0.003
<i>Comorbidity</i>					
Pulmonary embolism (%)	114 (1.0)	55 (1.1)	59 (1.0)	0.011	<0.001
Deep venous thrombosis (%)	292 (2.6)	196 (3.8)	96 (1.6)	0.139	0.015
Present of prosthetic heart valve (%)	364 (3.2)	292 (5.7)	72 (1.2)	0.249	0.01
Myocardial infarction (%)	3812 (33.7)	1281 (24.8)	2531 (41.2)	0.354	0.002
Heart failure (%)	3048 (27.0)	1692 (32.8)	1356 (22.1)	0.242	0.006
Atrial fibrillation (%)	2438 (21.6)	1630 (31.6)	808 (13.2)	0.454	0.002
Heart valve disease (%)	2743 (24.3)	1738 (33.7)	1005 (16.4)	0.408	0.005
Congenital heart disease (%)	293 (2.6)	197 (3.8)	96 (1.6)	0.14	0.002
Disorders of lipid metabolism (%)	9932 (87.9)	4629 (89.7)	5303 (86.4)	0.104	<0.001
Atherosclerosis (%)	3788 (33.5)	2064 (40.0)	1724 (28.1)	0.254	0.003

**Table 1** (continued)

	Overall (N = 11,298)	CABG (N = 5158)	PCI (N = 6140)	SMD Unadjusted	SMD MV-weighted
Intracerebral hemorrhage (%)	70 (0.6)	38 (0.7)	32 (0.5)	0.027	0.004
Gastrointestinal bleeding (%)	541 (4.8)	263 (5.1)	278 (4.5)	0.027	0.004
Other bleeding (%)	1360 (12.0)	681 (13.2)	679 (11.1)	0.066	0.003
Stroke or cardiovascular accident (%)	2524 (22.3)	1397 (27.1)	1127 (18.4)	0.209	0.007
High blood pressure (%)	9855 (87.2)	4630 (89.8)	5225 (85.1)	0.141	0.007
Diabetes mellitus (%)	4986 (44.1)	2498 (48.4)	2488 (40.5)	0.16	<0.001
Arrhythmias (%)	4689 (41.5)	2501 (48.5)	2188 (35.6)	0.263	0.003
Chronic kidney disease (%)	3452 (30.6)	1830 (35.5)	1622 (26.4)	0.197	0.002
Dementia (%)	49 (0.4)	12 (0.2)	37 (0.6)	0.057	0.002
Acute pulmonary heart disease (%)	683 (6.0)	467 (9.1)	216 (3.5)	0.23	0.009
Chronic pulmonary heart disease (%)	3027 (26.8)	1531 (29.7)	1496 (24.4)	0.12	0.003
Rheumatic disease (%)	287 (2.5)	128 (2.5)	159 (2.6)	0.007	0.001
Peptic ulcer disease (%)	196 (1.7)	95 (1.8)	101 (1.6)	0.015	0.002
Hemiplegia (%)	103 (0.9)	46 (0.9)	57 (0.9)	0.004	0.001
Neurologic disorder (%)	601 (5.3)	301 (5.8)	300 (4.9)	0.042	<0.001
Hypothyroidism (%)	1439 (12.7)	681 (13.2)	758 (12.3)	0.026	0.002
Malignant cancer (%)	851 (7.5)	393 (7.6)	458 (7.5)	0.006	0.006
Liver disease (%)	671 (5.9)	329 (6.4)	342 (5.6)	0.034	0.005
Human immunodeficiency virus (Hiv) (%)	1263 (11.2)	1042 (20.2)	221 (3.6)	0.53	0.009
Obesity (%)	3309 (29.3)	1639 (31.8)	1670 (27.2)	0.101	0.003
Weight loss (%)	577 (5.1)	300 (5.8)	277 (4.5)	0.059	0.006
Anemia (%)	436 (3.9)	226 (4.4)	210 (3.4)	0.05	0.001
Drug abuse (%)	1296 (11.5)	543 (10.5)	753 (12.3)	0.055	0.006
Psychosis (%)	133 (1.2)	64 (1.2)	69 (1.1)	0.011	0.003

Table 1 (continued)

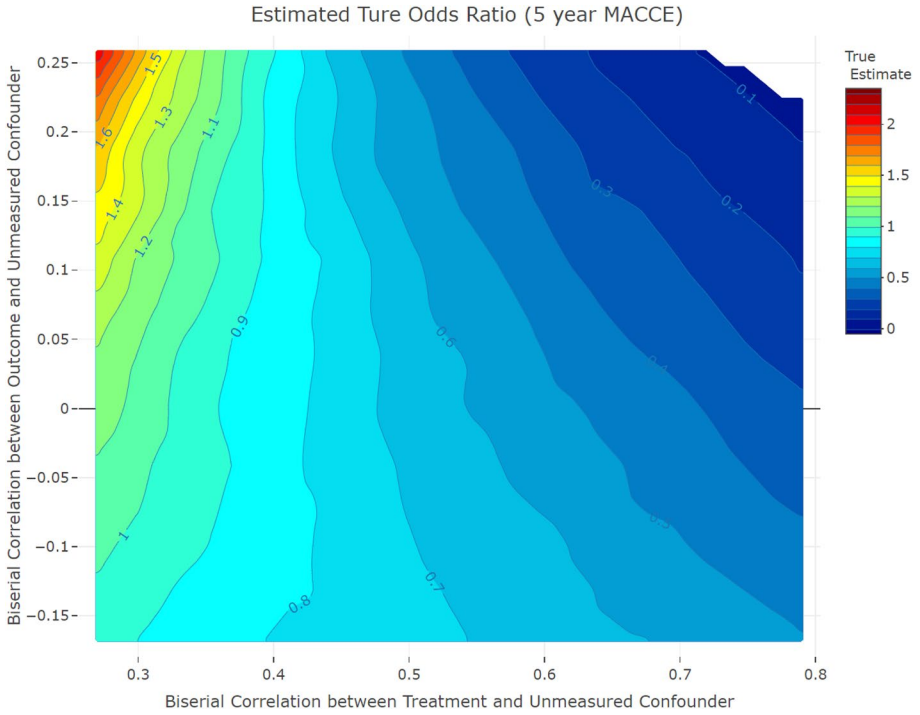
	Overall (N = 11,298)	CABG (N = 5158)	PCI (N = 6140)	SMD Unadjusted	SMD MV-weighted
Depression (%)	1716 (15.2)	757 (14.7)	959 (15.6)	0.026	0.004
<i>Medication history</i>					
Implantable cardioverter defibrillator (%)	155 (1.4)	52 (1.0)	103 (1.7)	0.058	<0.001
Angiotensin converting enzyme (Ace) inhibitor and angiotensin-receptor blockers (Arbs) (%)	7143 (63.2)	3456 (67.0)	3687 (60.0)	0.145	0.007
Aldosterone Antagonist (%)	273 (2.4)	136 (2.6)	137 (2.2)	0.026	0.002
Antiarrhythmic drugs (%)	262 (2.3)	138 (2.7)	124 (2.0)	0.043	0.002
Anticoagulants (%)	874 (7.7)	457 (8.9)	417 (6.8)	0.077	0.005
Antiplatelet drugs (%)	2162 (19.1)	676 (13.1)	1486 (24.2)	0.288	0.009
Beta blocker (%)	6983 (61.8)	3363 (65.2)	3620 (59.0)	0.129	0.001
Calcium channel blockers (%)	3051 (27.0)	1530 (29.7)	1521 (24.8)	0.11	0.003
Colchicine (%)	306 (2.7)	138 (2.7)	168 (2.7)	0.004	0.005
Digoxin (%)	377 (3.3)	189 (3.7)	188 (3.1)	0.033	0.003
Diuretics (%)	1849 (16.4)	952 (18.5)	897 (14.6)	0.104	0.004
Hydralazine (%)	506 (4.5)	257 (5.0)	249 (4.1)	0.045	<0.001
Nitrate (%)	1658 (14.7)	771 (14.9)	887 (14.4)	0.014	0.001
Statins (%)	7868 (69.6)	3809 (73.8)	4059 (66.1)	0.169	0.004

Mycardial infarction reported included MI diagnosed for only once at outpatient setting during baseline or non-primary diagnosis

**Table 2** Crude and adjusted treatment effects of CABG and PCI at the end of 1, 3, 5, 10 years

Year	Endpoint	Crude			Matching weight adjusted			L-table adjusted	
		OR (95% CI)	p value	E value	OR (95% CI)	p value	E value	OR (95% CI)	
1 years	MACCE	0.861 (0.789, 0.94)	0.001	1.367	0.713 (0.643, 0.789)	<0.001	1.651	0.424 (0.396, 0.517)	
	Death	0.998 (0.902, 1.103)	0.965	1.033	0.766 (0.68, 0.862)	<0.001	1.546		
	Stroke	1.263 (0.882, 1.817)	0.203	1.497	1.186 (0.78, 1.802)	0.424	1.4		
	MI	0.494 (0.409, 0.591)	<0.001	2.198	0.498 (0.404, 0.614)	<0.001	2.186		
	Repeat revascularization	0.494 (0.45, 0.543)	<0.001	2.198	0.457 (0.379, 0.55)	<0.001	2.321		
3 years	MACCE	0.849 (0.792, 0.91)	<0.001	1.39	0.701 (0.646, 0.761)	<0.001	1.676	0.436 (0.369, 0.527)	
	Death	0.912 (0.844, 0.986)	0.02	1.269	0.719 (0.656, 0.788)	<0.001	1.639		
	Stroke	1.277 (1.018, 1.605)	0.035	1.513	1.167 (0.897, 1.519)	0.249	1.375		
	MI	0.57 (0.492, 0.657)	<0.001	1.98	0.558 (0.473, 0.659)	<0.001	2.012		
	Repeat revascularization	0.466 (0.408, 0.53)	<0.001	2.29	0.486 (0.42, 0.562)	<0.001	2.224		
5 years	MACCE	0.913 (0.86, 0.97)	0.003	1.267	0.759 (0.707, 0.814)	<0.001	1.56	0.511 (0.451, 0.538)	
	Death	0.973 (0.912, 1.039)	0.418	1.132	0.785 (0.727, 0.847)	<0.001	1.51	0.457 (0.376, 0.468)	
	Stroke	1.159 (0.979, 1.372)	0.087	1.364	1.023 (0.842, 1.243)	0.818	1.119	0.694 (0.551, 0.873)	
	MI	0.583 (0.512, 0.662)	<0.001	1.947	0.57 (0.492, 0.661)	<0.001	1.98	0.358 (0.264, 0.368)	
	Repeat revascularization	0.48 (0.427, 0.537)	<0.001	2.243	0.497 (0.437, 0.565)	<0.001	2.189	0.304 (0.235, 0.344)	
10 years	MACCE	1.052 (0.997, 1.111)	0.064	1.188	0.893 (0.839, 0.95)	<0.001	1.306	0.505 (0.446, 0.582)	
	Death	1.1 (1.041, 1.163)	0.001	1.275	0.917 (0.86, 0.977)	0.008	1.259		
	Stroke	1.235 (1.086, 1.406)	0.001	1.463	1.111 (0.958, 1.289)	0.165	1.293		
	MI	0.667 (0.599, 0.741)	<0.001	1.749	0.667 (0.59, 0.753)	<0.001	1.749		
	Repeat revascularization	0.503 (0.455, 0.554)	<0.001	2.17	0.518 (0.464, 0.579)	<0.001	2.125		

MI myocardial infarction, MACCE major adverse cardiac and cerebrovascular events



**Fig. 4** Contour plot generated by L-table method illustrating estimated true effects corresponding to specified correlations

**Table 3** Sample of partial L-table where the true effect is identified (MACCE at 5 year,  $OR_{est}=0.759$ ,  $\rho_{TU}=0.586$ ,  $\rho_{YU}=0.055$ )

$\rho_{YU} \backslash \rho_{TU}$	<b>0.537</b>	<b>0.551</b>	<b>0.565</b>	<b>0.579</b>	<b>0.593</b>	<b>0.607</b>	<b>0.621</b>	<b>0.635</b>
<b>0.12</b>	0.528	0.506	0.484	0.459	0.43	0.404	0.38	0.362
<b>0.108</b>	0.54	0.518	0.496	0.471	0.442	0.415	0.391	0.373
<b>0.097</b>	0.553	0.53	0.508	0.483	0.453	0.426	0.404	0.385
<b>0.085</b>	0.56	0.542	0.521	0.495	0.464	0.44	0.417	0.397
<b>0.074</b>	0.566	0.549	0.532	0.507	0.48	0.454	0.43	0.409
<b>0.062</b>	0.576	0.558	0.541	0.52	0.492	0.465	0.441	0.423
<b>0.051</b>	0.589	0.571	0.552	0.53	0.502	0.475	0.452	0.435
<b>0.039</b>	0.602	0.582	0.561	0.541	0.513	0.486	0.463	0.446
<b>0.028</b>	0.609	0.588	0.567	0.549	0.522	0.495	0.474	0.457
<b>0.016</b>	0.607	0.586	0.568	0.553	0.525	0.503	0.488	0.468
<b>0.005</b>	0.605	0.588	0.573	0.556	0.532	0.517	0.502	0.482

\*  $\rho_{TU}$ , correlation coefficient between treatment and unmeasured confounder;  $\rho_{YU}$ , correlation coefficient between outcome and unmeasured confounder.

$\rho_{TU}$ , correlation coefficient between treatment and unmeasured confounder;  $\rho_{YU}$ , correlation coefficient between outcome and unmeasured confounder



**Table 4** Sensitivity analysis for variation of correlation inputs

Year	Endpoint	L-table adjusted $OR_{true}$ (95% CI)		
		$\rho_{TU}=0.625$	$\rho_{TU}=0.5$	$\rho_{TU}=0.7$
1 years	MACCE	0.424 (0.396, 0.517)	0.342 (0.333, 0.431)	0.456 (0.416, 0.592)
3 years	MACCE	0.436 (0.369, 0.527)	0.373 (0.349, 0.458)	0.459 (0.433, 0.575)
5 years	MACCE	0.511 (0.451, 0.538)	0.45 (0.397, 0.516)	0.517 (0.459, 0.607)
	Death	0.457 (0.376, 0.468)	0.379 (0.337, 0.416)	0.428 (0.397, 0.485)
	Stroke	0.694 (0.551, 0.873)	0.538 (0.45, 0.702)	0.824 (0.649, 1.035)
	MI	0.358 (0.264, 0.368)	0.258 (0.241, 0.326)	0.411 (0.283, 0.436)
	Repeat revascularization	0.304 (0.235, 0.344)	0.272 (0.213, 0.318)	0.329 (0.256, 0.412)
10 years	Death	0.505 (0.446, 0.582)	0.501 (0.421, 0.539)	0.632 (0.534, 0.652)

*MI* myocardial infarction, *MACCE* major adverse cardiac and cerebrovascular events

due to stroke at 5 years was no longer statistically significant (OR [95% CI] 0.824 [0.649, 1.035]) when we assign  $\rho_{TU}=0.7$  in simulation (Table 4).

## 4 Discussion

We introduce the L-table and illustrate how it can be used to make potential causal inferences using observational data. We have illustrated the application of the L-table in a real-world example, studying the safety and effectiveness of CABG vs. PCI for patients with ischemic heart disease at 1, 3, 5 and 10 years. We incorporated external information into the simulation with customizable parameters to locate the plausible range of estimated true effects. The foundation of the L-table is the correlation matrix for an ordinary linear estimate in simple or multivariate linear regression, but can be applied to most/many generalized linear models.

We found CABG is associated with better safety and effectiveness than PCI for patients with stable ischemic heart disease. After L-table adjustment, we found that the Odds Ratio shifted to smaller values, suggesting that unmeasured confounding decreased CABG's estimated advantage in standard analyses. We would have underestimated CABG's greater protective effect if disease severity were not accounted for.

In RCTs, the correlation between treatment and unmeasured confounders is approximately zero by design. In observational studies we found that the correlations between treatment selection and SYNTAX score are similar in a US-based study (Valle et al. 2019) (which we used in this study) and in a South Korean study (Kim et al. 2010), suggesting that clinicians from different countries share treatment preferences. The L-table adjusted effects vary depending on the simulation inputs. In this setting, sensitivity analysis may be more important than literal interpretation of standard errors. Therefore, users are advised to address the generalizability of the external data they refer to.

To choose between CABG and PCI is more complex in real life than in a RCT setting. Our results confirmed the findings from the RCTs, and validated the generalizability of the inference from the SYNTAX trials to our study. RCTs have great internal validity but may have poor external validity. The strategy we deployed in this study—extracting information (i.e., correlations) from external sources (i.e., RCTs), and integrating the information into

our observational data, enables us to ask the following question: do we see the same results in our system and population (in the real world) as found in RCTs?

Existing sensibility tests on unmeasured confounders primarily work with a two-by-two-by-two table setting. Although this dichotomous-variable setting allows the application of convenient mathematical relationships, it does not always generalize to broader settings. The L-table approach allows continuous measures in each component and by dichotomization or transformation facilitates generating a broad range of measure types. We have set up an infrastructure to move into a richer class of problems, and to help investigators broaden the application areas. In addition, conventional sensitivity analysis often only tells the users how much unmeasured confounding is needed to explain away the effects or statistical significance, while the L-table approach considers the plausible true effects. This can support cost effectiveness analysis and other practical decision making. Furthermore, the L-table approach exploits existing knowledge about unmeasured confounders to adjust the estimated effects. Future work includes modifying the L-table framework for the application of other analytical methods (i.e., time to event models, non-model-based randomization inference) or more complicated scenarios (i.e., reverse causality).

L-table offers a framework that is flexible to use. The multivariate normal data generating process provides a foundation for users to assign different functional forms. In addition, we assumed the unmeasured confounding in the simulation is a composite influence from all unmeasured confounders that play a role in the causal relationships. Consequently, the user needs only deduce a single proxy based on her knowledge of the unmeasured confounding. Finally, with a generalized linear model functional form, users can apply different analytical approaches to modify the simulation method accordingly.

Our simulation produces datasets with comparable properties as the balanced (PS-adjusted) empirical data, but with the addition of a simulated unmeasured confounder (not available in the empirical data). Users are encouraged to perform a propensity score adjustment on their observational data. This step balances the treatment and control group on measured confounders, and allows the simple trivariate normal setup in the simulation. In this study, we used matching weights adjusted real-world data to emphasize the causal effects inference on patients whose characteristics are such that each treatment has an equivalent chance to be chosen (i.e., we have equipoise).

A few limitations should be mentioned. Our proposed method is based on multiple assumptions, but the assumptions can be modified by the user when they adapt the L-table's framework. Second, the L-table is simulation based, so it comes without a closed-form solution. Further work is warranted to develop a mathematically supported theory. Third, the correlation between SYNTAX scores and outcomes are small. Although many studies have supported the utility of SYNTAX scores in the selection of revascularization strategies, contradictory findings have undermined the validity of such scores (He et al. 2020). Finally, the L-table adjusted true effects are sensitive to the input parameters. We advise users discuss the assumptions and limitations of their selected external sources and proxies.

## 5 Conclusion

We found CABG is associated with better clinical outcomes than PCI when treating patients with stable ischemic heart disease. Our results confirmed the findings from SYNTAX trials and validated the generalizability of the inference from the RCTs with our

observational data. The L-table, built with customizable parameters, adaptable models, and modifiable precision, provides investigators with a more plausible value of estimated true effects based on the influence of unmeasured confounders than is currently possible via other sensitivity techniques. Our methods lay out richer information and clearer directions than existing sensitivity tools, and can thus better support clinical decision-making. Finally, the L-table provides investigators with a deeper understanding of the likely treatment effects in the real world, and hence engenders greater confidence in study results based on both RCTs and observational data. We recommend using the L-table as a supplement to available sensitivity analyses.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10742-022-00282-y>.

**Acknowledgements** The authors thank Ernest Shen, PhD, and Mathew McHaffie, PhD, for critically reading the manuscript.

**Funding** LD was awarded for a statistical methodology research proposal by Division of Biostatistics Research at the Department of Research & Evaluation, Kaiser Permanente Southern California.

## Declarations


**Conflict of interest** The authors declare no conflict of interest.

## References

- Akima, H.: A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points. *ACM Trans. Math. Softw.* **4**, 148–164 (1978)
- Akima, H.: Algorithm 761: scattered-data surface fitting that has the accuracy of a cubic polynomial. *ACM Trans. Math. Softw.* **22**, 362–371 (1996)
- Brumback, B.A., Hernan, M.A., Haneuse, S.J., Robins, J.M.: Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. *Stat. Med.* **23**, 749–767 (2004)
- Chiba, Y.: Sensitivity analysis for unmeasured confounding of attributable fraction. *Epidemiology* **23**, 175–176 (2012)
- Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L.: Smoking and lung cancer: recent evidence and a discussion of some questions. *J. Natl. Cancer Inst.* **22**, 173–203 (1959)
- Cox, N.R.: Estimation of the correlation between a continuous and a discrete variable. *Biometrics* **30**, 171–178 (1974)
- Derose, S.F., Contreras, R., Coleman, K.J., Koebnick, C., Jacobsen, S.J.: Race and ethnicity data quality and imputation using U.S. Census data in an integrated health system: the Kaiser Permanente Southern California experience. *Med. Care Res. Rev.* **70**, 330–345 (2013)
- Duan, L.: Unmeasured confounder. GitHub repository. <https://github.com/leweiduan/UmeasuredConfounder.git> (2021). Accessed 30 Aug 2021
- Greenland, S.: Basic methods for sensitivity analysis of biases. *Int. J. Epidemiol.* **25**, 1107–1116 (1996)
- Greenland, S.: Useful methods for sensitivity analysis of observational studies. *Biometrics* **55**, 990–991 (1999)
- Greenland, S.: Multiple-bias modeling for analysis of observational data. *J. R. Stat. Soc. Ser. A* **168**, 267–306 (2005)
- Greenland, S.: Accounting for uncertainty about investigator bias: disclosure is informative. *J. Epidemiol. Community Health* **63**, 593–598 (2009)
- Greenland, S., Mansournia, M.A.: Penalization, bias reduction, and default priors in logistic and related categorical and survival regressions. *Stat. Med.* **34**, 3133–3143 (2015)
- He, Y.M., Shen, L., Ge, J.B.: Fallacies and possible remedies of the SYNTAX score. *J. Interv. Cardiol.* **2020**, 8822308 (2020)
- Hernan, M.A., Hernandez-Diaz, S., Robins, J.M.: Randomized trials analyzed as observational studies. *Ann. Intern. Med.* **159**, 560–562 (2013)

- Imbens GW, Rubin DB. Sensitivity analysis and bounds (Chapter 22). In: Causal Inference for Statistics, Social, and Biomedical Sciences, pp. 496–512. Cambridge University Press, New York (2015)
- Kappetein, A.P., Feldman, T.E., Mack, M.J., Morice, M.C., Holmes, D.R., Stahle, E., Dawkins, K.D., Mohr, F.W., Serruys, P.W., Colombo, A.: Comparison of coronary bypass surgery with drug-eluting stenting for the treatment of left main and/or three-vessel disease: 3-year follow-up of the SYNTAX trial. *Eur. Heart J.* **32**, 2125–2134 (2011)
- Kim, Y.H., Park, D.W., Kim, W.J., Lee, J.Y., Yun, S.C., Kang, S.J., Lee, S.W., Lee, C.W., Park, S.W., Park, S.J.: Validation of SYNTAX (Synergy between PCI with Taxus and Cardiac Surgery) score for prediction of outcomes after unprotected left main coronary revascularization. *JACC Cardiovasc. Interv.* **3**, 612–623 (2010)
- Li, L., Greene, T.: A weighting analogue to pair matching in propensity score analysis. *Int. J. Biostat.* **9**, 215–234 (2013)
- Lim, G.B.: Coronary artery disease: cost-effectiveness of CABG surgery versus PCI in complex CAD. *Nat. Rev. Cardiol.* **11**, 557 (2014)
- Maddala, G.S.: Limited-Dependent and Qualitative Variables in Econometrics. Cambridge University Press, Cambridge (1983)
- Mohr, F.W., Morice, M.C., Kappetein, A.P., Feldman, T.E., Stahle, E., Colombo, A., Mack, M.J., Holmes, D.R., Jr., Morel, M.A., Van Dyck, N., Houle, V.M., Dawkins, K.D., Serruys, P.W.: Coronary artery bypass graft surgery versus percutaneous coronary intervention in patients with three-vessel disease and left main coronary disease: 5-year follow-up of the randomised, clinical SYNTAX trial. *Lancet* **381**, 629–638 (2013)
- R core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2021)
- Rosenbaum, P.R.: Discussing hidden bias in observational studies. *Ann. Intern. Med.* **115**, 901–905 (1991)
- Rosenbaum, P.: Design sensitivity and efficiency in observational studies. *J. Am. Stat. Assoc.* **105**, 692–702 (2010)
- Rosenbaum, P.R., Rubin, D.B.: Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *J. R. Stat. Soc.* **45**, 212–218 (1983)
- Rosenbaum, P.R., Small, D.S.: An adaptive Mantel-Haenszel test for sensitivity analysis in observational studies. *Biometrics* **73**, 422–430 (2017)
- Rothman, K.J.: Six persistent research misconceptions. *J. Gen. Intern. Med.* **29**, 1060–1064 (2014)
- Schneeeweiss, S.: Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol. Drug Saf.* **15**, 291–303 (2006)
- Serruys, P.W., Morice, M.C., Kappetein, A.P., Colombo, A., Holmes, D.R., Mack, M.J., Stahle, E., Feldman, T.E., van den Brand, M., Bass, E.J., Van Dyck, N., Leadley, K., Dawkins, K.D., Mohr, F.W., Investigators, S.: Percutaneous coronary intervention versus coronary-artery bypass grafting for severe coronary artery disease. *N. Engl. J. Med.* **360**, 961–972 (2009)
- Thuijs, D., Kappetein, A.P., Serruys, P.W., Mohr, F.W., Morice, M.C., Mack, M.J., Holmes, D.R., Jr., Curzen, N., Davierwala, P., Noack, T., Milojevic, M., Dawkins, K.D., da Costa, B.R., Juni, P., Head, S.J., Investigators, S.E.S.: Percutaneous coronary intervention versus coronary artery bypass grafting in patients with three-vessel or left main coronary artery disease: 10-year follow-up of the multicentre randomised controlled SYNTAX trial. *Lancet* **394**, 1325–1334 (2019)
- Valle, J.A., Glorioso, T.J., Bricker, R., Baron, A.E., Armstrong, E.J., Bhatt, D.L., Rao, S.V., Plomondon, M.E., Serruys, P.W., Kappetein, A.P., Sabik, J.F., Dressler, O., Stone, G.W., Waldo, S.W.: Association of coronary anatomical complexity with clinical outcomes after percutaneous or surgical revascularization in the veterans affairs clinical assessment reporting and tracking program. *JAMA Cardiol.* **4**, 727–735 (2019)
- Vanderweele, T.J., Arah, O.A.: Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology* **22**, 42–52 (2011)
- VanderWeele, T.J., Ding, P.: Sensitivity analysis in observational research: introducing the E-value. *Ann. Intern. Med.* **167**, 268–274 (2017)

## Authors and Affiliations

Lewei Duan<sup>1</sup>  · Ming-Sum Lee<sup>2</sup> · Jason N. Doctor<sup>3</sup> · John L. Adams<sup>4</sup>

<sup>1</sup> Department of Research and Evaluation, Kaiser Permanente Southern California, 100 S Los Robles Ave, 5th Floor, Pasadena, CA 91101, USA

<sup>2</sup> Department of Cardiology, Kaiser Permanente Los Angeles Medical Center, Los Angeles, CA, USA

<sup>3</sup> Price School of Public Policy, University of Southern California, Los Angeles, CA, USA

<sup>4</sup> Kaiser Permanente Center for Effectiveness and Safety Research and the Bernard J. Tyson School of Medicine, Pasadena, CA, USA