

Richard Wyss*, Mark Lunt, M. Alan Brookhart, Robert J. Glynn and Til Stürmer

Reducing Bias Amplification in the Presence of Unmeasured Confounding through Out-of-Sample Estimation Strategies for the Disease Risk Score

Abstract: The prognostic score, or disease risk score (DRS), is a summary score that is used to control for confounding in non-experimental studies. While the DRS has been shown to effectively control for measured confounders, unmeasured confounding continues to be a fundamental obstacle in non-experimental research. Both theory and simulations have shown that in the presence of unmeasured confounding, controlling for variables that affect treatment (both instrumental variables and measured confounders) amplifies the bias caused by unmeasured confounders. In this paper, we use causal diagrams and path analysis to review and illustrate the process of bias amplification. We show that traditional estimation strategies for the DRS do not avoid bias amplification when controlling for predictors of treatment. We then discuss estimation strategies for the DRS that can potentially reduce bias amplification that is caused by controlling both instrumental variables and measured confounders. We show that under certain assumptions, estimating the DRS in populations outside the defined study cohort where treatment has not been introduced, or in outside populations with reduced treatment prevalence, can control for the confounding effects of measured confounders while at the same time reduce bias amplification.

Keywords: bias amplification, prognostic score, unmeasured confounding, path analysis, causal diagrams

DOI 10.1515/jci-2014-0009

1 Introduction

Measured and unmeasured confounding present challenges in non-experimental, e.g. pharmacoepidemiologic research. To control for large numbers of measured confounders, summary scores are increasingly used. The propensity score (PS), defined as the conditional probability of treatment given a set of measured covariates, has become the most widely used summary score for confounding control [1, 2]. An alternative summary score to the PS is the prognostic score, also known as the disease risk score (DRS) [3]. Unlike the PS which models covariate associations with treatment, the DRS models the probability or rate of disease occurrence absent of exposure. In a recent paper, Hansen [3] formalized the theoretical framework for the

*Corresponding author: **Richard Wyss**, Department of Epidemiology, UNC Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7435, USA, E-mail: wyss@email.unc.edu

Mark Lunt, Arthritis Research UK Epidemiology Unit, Centre for Musculoskeletal Research, Institute of Inflammation and Repair, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK, E-mail: Mark.Lunt@manchester.ac.uk

M. Alan Brookhart, Department of Epidemiology, UNC Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7435, USA, E-mail: mabrook@email.unc.edu

Robert J. Glynn, Department of Biostatistics, Harvard School of Public Health, Brigham & Women's Hospital, Boston, MA, USA, E-mail: rglynn@rics.bwh.harvard.edu

Til Stürmer, Department of Epidemiology, UNC Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7435, USA, E-mail: til.sturmer@post.harvard.edu

prognostic score or DRS. Formally, a DRS is defined as any scalar or multi-dimensional function that, when conditioned on, induces independence between measured covariates and the potential outcome under control (discussed further in Section 3) [3]. Although applications of the DRS have been limited compared to the PS, use of DRSs in medical studies has increased in recent years. A number of recent studies have demonstrated the application of DRSs for confounding control in both simulated and substantive data [3–9].

While both PSs and DRSs control for measured confounders, unmeasured confounding continues to be a fundamental obstacle in pharmacoepidemiology and non-experimental studies in general. In the presence of unmeasured confounding, it has been shown that controlling for variables that do not affect the outcome except through treatment (instrumental variables) amplifies bias caused by unmeasured confounders [10–15]. Pearl [12] further explains that bias amplification is not just a function of controlling for instruments but also occurs when controlling for any variable that affects treatment, including measured confounders. Controlling for measured confounders, however, removes confounding bias due to the measured confounders in addition to increasing bias caused by unmeasured confounders.

Given the potential for bias amplification, PS and DRS models that exclude instrumental variables are desirable in terms of reducing bias caused by unmeasured confounders. Because bias amplification is also a function of controlling for measured confounders, Pearl [12] suggests that researchers should consider the cost when controlling for measured confounders that have a strong effect on treatment but only a weak effect on the outcome (near instruments). For studies involving large numbers of covariates, however, identifying instrumental variables and evaluating the cost of controlling for near instruments can be challenging. Pharmacoepidemiologic and medical studies utilizing automated databases often involve large numbers of potential covariates that have not been selected with a specific research question in mind and where a multitude of factors other than the prognosis strongly influence treatment decisions (e.g. marketing, formularies, and physician preference) [16]. In these settings, reducing bias amplification through automated or knowledge driven variable selection strategies can be difficult.

In this paper, we discuss ways in which researchers can estimate DRSs to potentially reduce bias amplification in situations where it is difficult to identify instrumental variables or evaluate the cost of controlling for near instruments. In Section 2, we use causal diagrams and path analysis to review the process of bias amplification. We show how bias amplification results from controlling indirect correlations that are induced between predictors of treatment and unmeasured confounders when conditioning on treatment. In Section 3, we review the balancing properties of DRSs and discuss why traditional estimation strategies for the DRS cause bias amplification in the presence of unmeasured confounders. In Section 4, we then discuss alternative out-of-sample estimation strategies for the DRS within large medical claims databases. We show that under certain assumptions, researchers can reduce bias amplification by estimating the DRS within historical data prior to treatment introduction or by estimating the DRS in populations with reduced treatment prevalence. In Section 5, we compare DRS estimation strategies when comparing the COX-2 inhibitor celecoxib to non-selective NSAIDs in reducing gastrointestinal complications. In Section 6, we conclude and discuss limitations of the proposed estimation strategies for the DRS.

2 Bias amplification

2.1 Setup and notation

For illustrative purposes, consider the causal diagram described in Figure 1 where the nodes represent random variables and the arrows represent causal effects. Figure 1 describes a causal structure consisting of a treatment (T), an instrumental variable (Z), an unmeasured confounder (U), and an outcome (Y). Causal associations between variables are shown by paths in which all arrows point forward. For example, the path $U \rightarrow Y$ represents the direct causal effect of U on Y and the path $U \rightarrow T \rightarrow Y$ represents the indirect

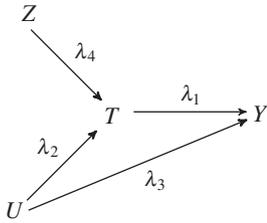


Figure 1 Causal diagram consisting of a treatment, T , an instrumental variable, Z , an unmeasured confounder, U , and an outcome, Y

causal effect of U on Y . For simplicity, we will assume that T is a linear function in Z and U and that Y is linear in U and T .

Under assumptions of linearity, Wright [17] developed a method of path analysis to determine the correlation between two variables. This method initially labeled the arrows between variables with correlation coefficients, since they are the same in both directions. However, to assess causal effects in terms of how much one variable will change when another is changed by a given amount, regression coefficients need to be used.

Lunt et al. [18] and Pearl [19] explain that Wright's method can be modified slightly to determine the relation between two variables in terms of regression coefficients. It has been shown that the value of a path between two variables is the product of the individual segments in the path and that the overall association between two variables is determined by the sum of all paths between the variables [17–19]. So in this case, if we were to perform a regression of Y on U (a hypothetical regression as U is unobserved), the coefficient for U would be equal to $\lambda_3 + \lambda_2\lambda_1$, where λ_3 represents the value of the direct causal effect of U on Y , and $\lambda_2\lambda_1$ represents the value of the indirect causal effect of U on Y through T (Figure 1).

If we wished to calculate the direct effect of U on Y , we would need to block the indirect path $U \rightarrow T \rightarrow Y$. Paths can be blocked either by controlling for a variable on the path or by having a “collider” on the path. For example, the path $U \rightarrow T \leftarrow Z$ is blocked by the collider at T . If we were to regress Y on U and T , the coefficient for U would be λ_3 .

Regression coefficients can also represent non-causal effects. For example, in Figure 1, we could regress T on U and the regression coefficient would be λ_2 . However, we could also regress U on T . It is important to note that this regression coefficient will not be equal to λ_2 , since its units will be whatever units U is measured in, whereas λ_2 represents a change in the prevalence of treatment when U changes by one unit. For this reason, we let the regression coefficient $r(\lambda_2)$ represent the value of the path labeled λ_2 , but in the reverse direction.

Causal diagrams enable us to identify and control for confounding effects. For example, in Figure 1 there are two paths from T to Y : $T \rightarrow Y$ and $T \leftarrow U \rightarrow Y$. The second path does not represent a causal effect since all of the arrows do not point in the forward direction. The coefficient $\lambda_1 + r(\lambda_2)\lambda_3$ represents the effect of T on Y through both of these paths. This coefficient does not represent the causal effect of T on Y , since it is confounded by U . If we block the path $T \leftarrow U \rightarrow Y$ by controlling for U (e.g. including it in the regression equation), only the direct causal path from T to Y is left open, and the regression coefficient for T will represent the causal effect. Pearl [12, 19] and Lunt et al. [18] provide detailed discussions on path analysis and additional examples of Wright's method in a structural equation setting.

2.2 Bias amplification from a causal diagram perspective

Figure 1 illustrates that the instrumental variable, Z , and the unmeasured confounder, U , are marginally independent since the path between Z and U is blocked by the collider, T . Once we condition on T , which we do when we estimate the treatment effect, we induce a correlation between Z and U which is represented by the dashed arrows between Z and U in Figure 2(a). The dashed lines illustrate that this association is

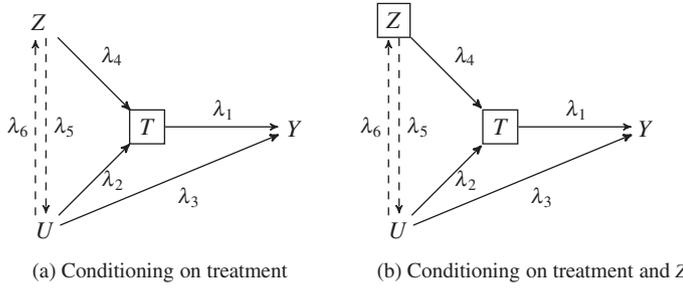


Figure 2 (a) Conditioning on T induces an association between Z and U . The dashed lines between Z and U illustrate that this association is indirect and non-causal. The parallel arrows illustrate that this induced association represents two pathways. (b) After conditioning on T and Z , all pathways from T to Y that occur through the instrument, Z , are blocked

induced and non-causal. The two arrows between Z and U are used to illustrate that this induced association represents two pathways: (1) the indirect path in the direction from Z to U ($Z \rightarrow U$) and (2) the indirect path in the direction from U to Z ($U \rightarrow Z$). The value of the path $Z \rightarrow U$ is represented by the term λ_5 and can be interpreted as the regression coefficient for Z in a regression of U on Z and T . Similarly, the value of the path $U \rightarrow Z$ is represented by the term λ_6 and is interpreted as the regression coefficient of U in a regression of Z on U and T .

There are now four paths from treatment, T , to the outcome, Y : (1) the direct path $T \rightarrow Y$, (2) the indirect path $T \leftarrow U \rightarrow Y$ (confounding by U), (3) the indirect path $T \leftarrow Z \rightarrow U \rightarrow Y$ (confounding by U and Z , created by the correlation between U and Z that occurs when conditioning on T), and (4) the indirect path $T \leftarrow U \rightarrow Z \rightarrow U \rightarrow Y$ (which also represents confounding by U and Z , created by the correlation between U and Z after conditioning on T). Were we to regress Y on T , the coefficient for T would represent the sum effect of these four paths. Formally,

$$\frac{\partial}{\partial t} E[Y|T = t] = \frac{\partial}{\partial t} \sum_z \sum_u E[Y|t, z, u] p(u|z, t) p(z|t) \quad (1)$$

$$= \lambda_1 + r(\lambda_2)\lambda_3 + \lambda_5\lambda_3 \frac{\partial}{\partial t} E[Z|t] \quad (2)$$

$$= \lambda_1 + r(\lambda_2)\lambda_3 + r(\lambda_4)\lambda_5\lambda_3 + r(\lambda_2)\lambda_6\lambda_5\lambda_3. \quad (3)$$

Each of the terms in eq. (3) represents one of the four pathways from T to Y when we regress Y on T . The first term, λ_1 , is the direct effect of T on Y and represents the unconfounded treatment effect. The second term in eq. (3), $r(\lambda_2)\lambda_3$, represents the value of the path $T \leftarrow U \rightarrow Y$. The third term, $r(\lambda_4)\lambda_5\lambda_3$, represents the value of the path $T \leftarrow Z \rightarrow U \rightarrow Y$. The last term, $r(\lambda_2)\lambda_6\lambda_5\lambda_3$, represents the value of the path $T \leftarrow U \rightarrow Z \rightarrow U \rightarrow Y$. The sum of the last three terms in eq. (3) represents the bias in the treatment effect when not controlling for the instrument, Z , and can be expressed as

$$B_0 = r(\lambda_2)\lambda_3 + r(\lambda_4)\lambda_5\lambda_3 + r(\lambda_2)\lambda_6\lambda_5\lambda_3. \quad (4)$$

After controlling for Z , all of the pathways between T and Y that occur through Z are blocked (Figure 2 (b)). There are now only two paths between treatment and the outcome: (1) the direct path $T \rightarrow Y$ and (2) the confounding path $T \leftarrow U \rightarrow Y$. A regression of Y on T and Z can then be calculated as the sum of the values of these two pathways.

$$\frac{\partial}{\partial t} E[Y|T = t, Z = z] = \frac{\partial}{\partial t} \sum_u E[Y|t, z, u] p(u|t, z) \quad (5)$$

$$= \lambda_1 + \lambda_3 \frac{\partial}{\partial t} E[U|t, z] \quad (6)$$

$$= \lambda_1 + r(\lambda_2)\lambda_3. \quad (7)$$

In eq. (7), λ_1 represents the value of the direct effect of T on Y and the term $r(\lambda_2)\lambda_3$ represents the value of the indirect path $T \leftarrow U \rightarrow Y$. The bias in the treatment effect after controlling for the instrument, Z , can then be expressed as

$$B_z = r(\lambda_2)\lambda_3. \quad (8)$$

Pearl [12] shows that $B_0 = (1 - \rho_{ZT}^2)B_z$, where ρ_{ZT} is the correlation between Z and T . Since $0 \leq \rho_{ZT}^2 \leq 1$, $B_0 \leq B_z$ with equality only if $\rho_{ZT} = 0$. In other words, not controlling for the instrument, Z , reduces the overall bias in the treatment effect that is caused by the unmeasured confounder, U . When we do not condition on Z , we allow for an induced correlation to occur between Z and U upon conditioning on T . This induced correlation results in additional confounding (an induced confounding) that is of smaller magnitude and in the opposite direction than the confounding caused by U alone. Therefore, the overall confounding in the treatment effect is less than the confounding caused only by U . Controlling for Z eliminates the induced confounding and returns the total confounding to the value it would have been in the absence of the instrument.

Pearl [12] provides an intuitive way of describing this effect which is discussed by Myers et al. [14]. According to Pearl, not conditioning on an instrumental variable allows the instrument to account for part of the change or variation in the treatment variable. This reduces the amount of variation in the treatment variable that is explained by the unmeasured confounder, thereby reducing the total amount of residual confounding caused by the unmeasured confounder.

While we have described bias amplification in terms of controlling for an instrumental variable, the same arguments apply when controlling for any variable affecting treatment (e.g. measured confounders). Controlling for a measured confounder, X , will remove the confounding due to X , but also eliminate the induced reduction in confounding by U . This may increase or decrease the overall confounding, depending on the relative strengths of these two confounding effects. For a detailed discussion on bias amplification, we refer the reader to Pearl [12]. The purpose for this discussion is to emphasize and illustrate that bias amplification results from controlling induced correlations that occur between instrumental variables (or measured confounders) and unmeasured confounders upon conditioning on treatment. Viewing bias amplification from this perspective helps to discuss ways in which DRSs can be estimated to avoid controlling these induced correlations while simultaneously controlling for the direct effects of measured confounders on the outcome.

3 Balancing properties of DRSs and bias amplification when conditioning on $E[Y_0|X]$ vs $E[Y|X, T = 0]$

3.1 Covariate balance when conditioning on prognostic scores

In medical research investigators are often interested in comparing a treatment to a comparator therapy or no treatment. Following Rubin's description of the counterfactual framework [20, 21], let Y_1 represent the potential response had the individual received treatment and Y_0 the potential response had the individual remained untreated. In practice, only one of the potential outcomes is observed. Let Y represent the observed outcome and T a dichotomous treatment. Further, let X represent a set of measured baseline covariates.

Hansen [3] shows that a prognostic score, or DRS, acts as a "prognostic balancing" score in that measured covariates are independent of the potential outcome under no treatment upon conditioning on the DRS. Formally, a function of X , $\psi(X)$, is a DRS if it satisfies $Y_0 \perp X | \psi(X)$, where \perp denotes the independence of random variables and $|$ denotes conditional on. Prognostic balance differs from "propensity balance" since conditioning on the PS renders covariates independent of treatment assignment, not the outcome. Hansen explains that if the conditional expectation of Y_0 given X follows a generalized linear model, then one possible

DRS is the function $E[Y_0|X]$. The DRS can also be a multi-dimensional function. For example, if $Y_0|X$ follows a linear model where the variance of Y_0 is nonconstant in X , then the function $E[Y_0|X]$ and the variance function together constitute a DRS [3]. Hansen further shows that under the assumption $Y_0 \perp T|X$ (i.e. treatment assignment is weakly ignorable or no unmeasured confounding), subclassification on a DRS, $\psi(X)$, is sufficient to satisfy $Y_0 \perp T|\psi(X)$ and allow for identification of the treatment effect in the treated.

3.2 Bias amplification when conditioning on $E[Y_0|X]$

Consider the causal scenario illustrated in Figure 3(a) where T represents a treatment, X a measured confounder, U an unmeasured confounder, and Y the outcome. For simplicity, we will keep the same linearity and distributional assumptions discussed previously when describing Figure 1 with the exception that we will relax the linearity assumption between X and Y and assume

$$E[Y|T = t, U = u, X = x] = \lambda_1 t + \lambda_3 u + f(x), \quad (9)$$

where $f(x)$ represents a non-linear function in X (e.g. $f(x) = x + x^2$). The direct effect of X on Y can then be represented by $f'(x) = \frac{\partial}{\partial x} f(x)$ as shown in Figure 3(a).

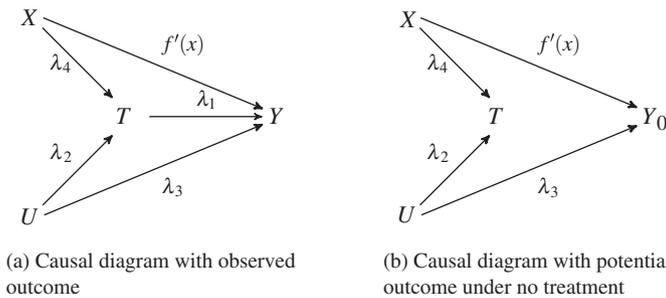


Figure 3 (a) Causal diagram consisting of a dichotomous treatment, T , a measured confounder, X , an unmeasured confounder, U , and an outcome, Y . (b) Causal diagram illustrating the relation between covariates in Figure 3(a) with the potential outcome under no treatment, Y_0

The reason for not assuming linearity between Y and X is because if the DRS is a linear function of only one covariate, X , then each DRS value could be mapped to a unique value of X and conditioning on the DRS would essentially be equivalent to conditioning on X itself. When the DRS is non-linear in X or a function of two or more covariates, however, this is generally not the case. While the previous discussion on bias amplification assumed a linear structural equation framework, Pearl [12, 22] extends the discussion of bias amplification to non-linear models and shows that similar arguments for describing bias amplification can apply in many non-linear settings.

When discussing confounding control through DRSs, it is helpful to illustrate prognostic balance using DAGs that include the potential outcome under no treatment, Y_0 . Richardson and Robins [23] have formalized the unification of DAGs with counterfactual theory and potential outcomes. Figure 3(b) shows the causal relations between the covariates X , U , and T with Y_0 . Because Y_0 is defined as the potential outcome had everyone remained untreated, the observed treatment variable, T , does not affect Y_0 [23]. Because we have assumed that Y is linear in T , the direct effects of X and U on Y_0 are the same as those for the observed outcome, Y (Figure 3).

For this causal scenario, one possible DRS for the measured covariate X is the function $E[Y_0|X]$. It is straightforward to show that $E[Y_0|X]$ is equivalent to modeling the direct relationship between X and Y_0 which, in this case, is represented by the function $f(X)$ in eq. (9). Therefore, conditioning on $E[Y_0|X]$ blocks the direct pathway between X and Y_0 as illustrated in Figure 4(a): it can be thought of as blocking the arrow labeled $f'(X)$ from the figure, resulting in Y_0 being independent of X (i.e. $Y_0 \perp X|E[Y_0|X]$).

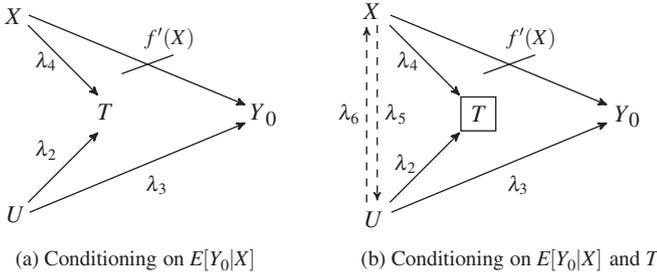


Figure 4 (a) Conditioning on $E[Y_0|X]$ blocks the direct path from X to Y_0 which results in Y_0 being independent of X (i.e. $Y_0 \perp X | E[Y_0|X]$). (b) After conditioning on both $E[Y_0|X]$ and T , X becomes associated with Y_0 through an induced association between X and U

If we condition on T , we again induce a correlation between U and X that is not modeled in the DRS function, $E[Y_0|X]$. So, if we regress Y_0 on T and $E[Y_0|X]$, there are three pathways between T and Y_0 : (1) the path $T \leftarrow U \rightarrow Y_0$, (2) the path $T \leftarrow X \rightarrow U \rightarrow Y_0$, and (3) the path $T \leftarrow U \rightarrow X \rightarrow U \rightarrow Y_0$ (Figure 4(b)). Paths (2) and (3) partially counter the effect of confounding by path (1) as discussed previously. Therefore, the bias in the estimated treatment effect when conditioning on $E[Y_0|X]$ is reduced compared to the bias that would result had we controlled for X directly since conditioning on X blocks paths (2) and (3) while conditioning on $E[Y_0|X]$ does not (Figure 4(b)).

3.3 Bias amplification when conditioning on $E[Y|X, T = 0]$

In practice, Y_0 is only observed for the untreated population, and the DRS must be estimated indirectly using the observed outcome Y . This has primarily been done in one of two ways. The first is to fit a regression model to the untreated within the study cohort and then use this model to predict the disease risk for all individuals within the full cohort. The second is to fit a regression model to the full cohort (i.e. both treated and untreated) as a function of baseline covariates and treatment and then estimate the disease risk for each individual after setting treatment status to untreated [3, 5–7, 9, 24].

Both of these estimation strategies attempt to estimate $E[Y_0|X]$ indirectly by estimating the function $E[Y|X, T = 0]$. If there were no unmeasured confounding, then it is easy to show that $E[Y|X, T = 0] = E[Y_0|X]$. In the presence of unmeasured confounding, however, $E[Y|X, T = 0] \neq E[Y_0|X]$ and $E[Y|X, T = 0]$ does not have the properties of a prognostic score (or DRS) as defined by Hansen [3]. Further, in the presence of unmeasured confounding, conditioning on $E[Y|X, T = 0]$ results in bias amplification. This is because the function $E[Y|X, T = 0]$ conditions (or restricts) on $T = 0$ and, therefore, the regression coefficient for X in this function will not only represent the direct effect of X on Y but also the indirect correlation that is induced between X and Y through U when conditioning on $T = 0$. Consequently, conditioning on $E[Y|X, T = 0]$ not only controls for the confounding due to X but also controls for the induced correlation between X and Y through U and, therefore, removes the reduction in confounding in U that is caused by this induced correlation. Hence, the total confounding effect in this case is B_X (eq. 8) which is greater than the overall confounding had we conditioned on $E[Y_0|X]$.

3.4 Simulation illustrating bias amplification when conditioning on $E[Y_0|X]$ vs $E[Y|X, T = 0]$

For illustrative purposes, we simulated a causal structure consisting of a dichotomous treatment, T , a standard normal measured confounder, X , a standard normal unmeasured confounder, U , and a normally distributed outcome, Y . The conditional expectation of treatment and outcome were simulated according to

eqs (10) and (11). We varied the strength of the effect of the measured confounder, X , on treatment (α_1) while holding the other parameters constant.

$$\text{logit}(E[T|x, u]) = \alpha_1 x + 0.7u \quad (10)$$

$$E[Y|x, u, t] = 0.7x + 0.7x^2 + 0.7u + 0.7t \quad (11)$$

The parameter values in eqs (10) and (11) were chosen simply for illustrative purposes, and other values could be used. In addition to the scenarios described in eqs (10) and (11) where U acts as an unmeasured confounder, we also conducted simulations where U had no effect on T or Y (no unmeasured confounding) to illustrate that traditional estimation strategies for the DRS can successfully control for measured confounders and result in unbiased effect estimates in the absence of unmeasured confounding.

We estimated the treatment effect after conditioning on $E[Y_0|X]$ and again after conditioning on $E[Y|X, T = 0]$. Because the data were simulated, $E[Y_0|X]$ was known by design. We estimated $E[Y|X, T = 0]$ by fitting an outcome model within the untreated cohort and then used this model to predict disease risk for each individual within the full cohort. For comparison, we also estimated the treatment effect after adjusting for X directly within a traditional outcome regression model.

Results in Table 1 show the bias in the unadjusted treatment effect as well as the bias after adjusting for X directly in an outcome regression model, or adjusting for X by conditioning on $E[Y_0|X]$ or $E[Y|X, T = 0]$. When there is no unmeasured confounding and X does not affect treatment ($\text{OR} = 1.0$), both the unadjusted and the adjusted effect estimates are unbiased as expected since there is no confounding due to either U or X . When there is no unmeasured confounding and X has an effect on treatment (X is a measured confounder), adjusting for X directly in an outcome regression model, or conditioning on either $E[Y|X, T = 0]$ or $E[Y_0|X]$, successfully controls for the confounding caused by X , resulting in approximately unbiased effect estimates (Table 1).

When U acts as an unmeasured confounder and X has no effect on T ($\text{OR} = 1$), the unadjusted and adjusted bias is 0.44. This bias represents the magnitude of confounding due to the unmeasured confounder, U . As the strength of the effect of X on T increases, the unadjusted bias also increases as expected (Table 1). When we condition on $E[Y_0|X]$, we control for the confounding that is caused by X without controlling the reduction in confounding in U that is due to the induced correlation between X and U as discussed previously. This reduction in confounding that is caused by the induced correlation between X and U is illustrated by the reduction in bias for $E[Y_0|X]$ when X affects treatment (Table 1).

When U acts as an unmeasured confounder and we condition on $E[Y|X, T = 0]$ or control for X directly in an outcome regression model, we remove the confounding that is caused by X , but we also control the reduction in confounding in U that is due to the indirect correlation between X and U as discussed

Table 1 Simulation results

Scenario	$\exp(\alpha_1)^b$	Absolute bias in the estimated treatment effect ^a			
		Unadjusted	Outcome regression	$E[Y X, T = 0]$	$E[Y_0 X]$
No unmeasured confounding					
	OR = 1.0	0.00	0.00	0.00	0.00
	OR = 3.0	0.62	0.00	0.00	0.00
	OR = 5.0	0.77	0.00	0.00	0.00
With unmeasured confounding					
	OR = 1.0	0.44	0.44	0.44	0.44
	OR = 3.0	0.95	0.44	0.44	0.39
	OR = 5.0	1.06	0.44	0.44	0.35

Notes: ^aBias after adjusting for X directly in a traditional outcome regression model, or adjusting for X by conditioning on $E[Y|X, T = 0]$ or $E[Y_0|X]$; ^bStrength of the effect of X on T .

previously. Therefore, the overall confounding in the estimated treatment effect is the same had this indirect correlation not existed at all (i.e. X did not affect T to begin with). While bias amplification is the common term used to describe this event, this process can be thought of as a bias reduction that occurs when not controlling indirect correlations that are induced between predictors of treatment and unmeasured confounders.

We use this example to emphasize two aspects of bias amplification: (1) IVs and measured confounders can reduce the magnitude of bias caused by unmeasured confounders when the induced correlations between predictors of treatment and unmeasured confounders are not controlled and (2) prognostic scores (or DRSs) can control for the confounding effects of measured confounders while at the same time reduce bias amplification.

4 Estimation strategies for the DRS to reduce bias amplification

4.1 Challenges with same-sample estimation

Under the assumption of no unmeasured confounding, the traditional DRS estimation strategies discussed previously can result in accurate estimates for the DRS. Even in settings with no unmeasured confounding, however, implementing these strategies in practice can be challenging due to limitations when modeling the DRS within the study cohort. Fitting the DRS to the full cohort benefits from increased sample size, but can be particularly sensitive to model misspecification due to the additional assumptions required for correctly modeling the relation between the treatment and outcome [3]. Hansen [3] explains that even small misspecifications in the specified DRS model can result in the treatment effect being associated with the estimated scores. This non-ancillarity in the estimated DRSs can introduce additional bias when used for confounding control [3, 9].

Fitting the DRS only within the untreated cohort, however, increases the potential for overfitting the model which can result in misspecified DRSs [3, 7]. In particular, if the DRS is estimated from the untreated or control group within the study data, this estimated DRS model will be overfit in this referent group, leading to over-estimation of risk in the control group for higher risk patients and under-estimation of risk in the control group for lower risk patients [3, 7]. Since the DRS is often and appropriately used for stratified treatment comparisons, such within study estimation can introduce bias in these comparisons. While bias due to overfitting the DRS model may be small in some settings (e.g. the bias due to overfitting the DRS model in the simulations described in Section 3.4 was very small which is likely due to the very large sample size used in the simulations), Hansen [3] describes scenarios and gives an example involving smaller sample sizes where overfitting the DRS model can potentially be problematic. Because factors affecting disease risk are likely to be constant over time and across populations, recent strategies that use out-of-sample estimation have been proposed to potentially overcome these shortcomings.

Both Hansen [3] and Glynn et al. [7] have proposed that disease risk can be accurately estimated from either a separate population, or the same population but with historical data from a period prior to the current study period. Hansen explains that estimating the DRS within an alternate sample of controls can potentially avoid the complications of overfitting that can occur when using same-sample estimation. Glynn et al. [7] suggest that estimating the DRS within historical data prior to treatment introduction can be particularly advantageous in pharmacoepidemiologic studies that use large administrative healthcare databases to evaluate newly introduced treatments and evolving drug therapies. Estimation with historical data from the same data collection process increases the comparability of subject surveillance and variable definitions between the DRS model development and application samples and limits overfitting associated with the application of a prediction model to the same data used in its development.

4.2 Estimating disease risk within historical data prior to treatment introduction

A potential advantage of estimating the DRS within historical data prior to treatment introduction that has not been discussed is the potential to reduce bias amplification in the presence of unmeasured confounders. Recall that bias amplification results from controlling indirect associations between predictors of treatment and unmeasured confounders that are induced upon conditioning on treatment. When using historical data prior to treatment introduction to estimate the DRS, we can avoid controlling these induced associations since we avoid having to condition on treatment when estimating the DRS.

For example, consider the same causal structure discussed previously with the exception that treatment has not been introduced (Figure 5(a)). Assuming the effects of covariates on the observed outcome are constant across populations, it is straightforward to show that $E[Y_h|X]$ is equal to $E[Y_0|X]$, where Y_h represents the observed outcome within the historical population prior to treatment introduction. Therefore, conditioning on $E[Y_h|X]$ within the original study cohort where treatment is present satisfies $Y_0 \perp X|E[Y_h|X]$ by blocking the direct effect of X on Y_0 without controlling the indirect association between X and Y_0 through U (Figure 5(b)).

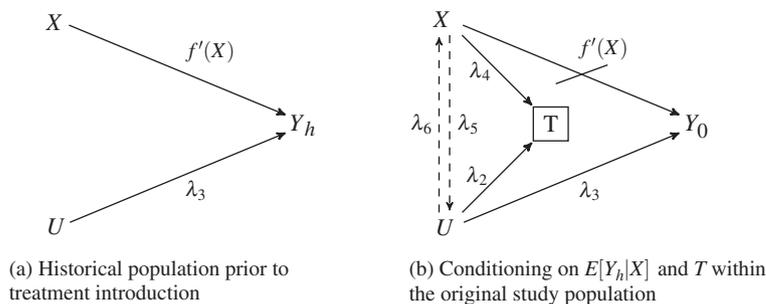


Figure 5 (a) Causal diagram representing a historical population prior to the introduction of treatment. (b) Population after the introduction of treatment. Assuming the effects of covariates on the outcome are constant across populations, conditioning on $E[Y_h|X]$ blocks the direct effect of X on Y_0 without controlling the indirect association between X and Y_0 that occurs through U

4.3 Estimating disease risk in populations with reduced treatment prevalence

Although the theoretical aspects of using historical data to estimate the DRS are promising, the implementation of this strategy is limited to situations where historical data are available. For studies conducted within administrative healthcare databases, this strategy can further be challenging when covariate assessments and coding practices change over time. To address these complications we propose an alternative DRS estimation strategy for studies conducted within large administrative databases: estimate the DRS in populations where the treatment prevalence is reduced compared to the original study population.

In pharmacoepidemiologic and medical studies, many restrictions are often placed on study participation to increase the comparability of treatment groups. For example, study participation is often restricted to new-users of a given treatment (or treatments) after a specified washout period to reduce possible biases caused by healthy users [25]. The new-user design and other restriction criteria are necessary to reduce bias when using non-experimental study designs to evaluate medical treatments. However, many restrictions placed on study participation when estimating the treatment effect may be unnecessarily restrictive when estimating the DRS. If the effects of covariates on disease risk are stable across populations, then individuals from outside the defined study cohort can potentially be used when estimating the DRS. Including external data increases the volume of outcome events which can be advantageous when attempting to model rare outcomes as a complex function of large sets of covariates. Further, including

external data when estimating the DRS can potentially allow researchers to reduce bias amplification by lowering the prevalence of treatment in the population where the DRS is estimated.

The strength of induced correlations between predictors of treatment and unmeasured confounders that occur when conditioning on treatment is a function of the prevalence of treatment within the population. Reducing the prevalence of treatment reduces the strength of these induced correlations (see Appendix for a detailed explanation). For example, consider two populations with similar causal structures with the exception that the prevalence of treatment is lower in one of the populations (Figure 6(a) and 6(b)). Let Y_r and T_r represent the observed outcome and treatment in the population with reduced treatment prevalence. The induced correlation between X and U that occurs when conditioning on treatment will be weaker within the population with reduced treatment prevalence compared to the original cohort (Figure 6(a) and 6(b)).

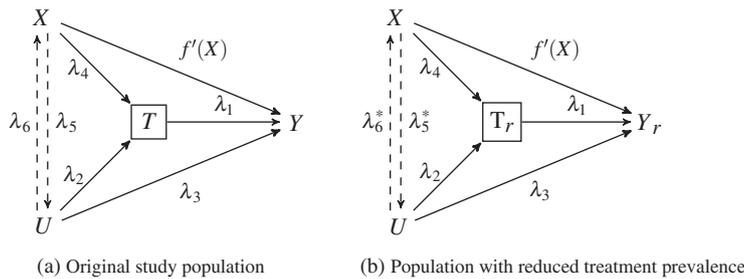


Figure 6 (a) Original study cohort where the treatment effect is estimated. (b) Population with reduced treatment prevalence where the DRS is estimated. The induced association between X and U after conditioning on T_r is weaker than the induced association between X and U within the original study cohort after conditioning on T , and the regression coefficients λ_5^* and λ_6^* in Figure 6(b) are smaller in magnitude than the regression coefficients λ_5 and λ_6 in Figure 6(a)

If we estimate the DRS within the population with reduced treatment prevalence, we will be estimating a function that is equal to $E[Y_r|X, T_r = 0]$. The regression coefficient for X in this function represents the direct effect of X on Y_r as well as the indirect association between X and Y_r that occurs through U when conditioning on $T_r = 0$. This coefficient correctly represents the direct effect of X on Y within the original study cohort since the effect of X on the outcome is constant across populations. However, this coefficient also underrepresents the strength of the induced association that X has with Y through U in the original population since the induced correlation between X and U is weaker in the population with reduced treatment prevalence where the DRS was estimated. Therefore, conditioning on $E[Y_r|X, T_r = 0]$ within the original study population controls for the confounding caused by X , but only partially controls the induced correlation between X and U within the original study cohort. This correlation, in turn, reduces the magnitude of confounding caused by U as discussed previously.

4.4 Simulation study: a simple illustrative example

To illustrate, we simulated the same population described in eqs (10) and (11) which has a baseline treatment prevalence of 50%. In this example, we held the effect of X on T constant at an OR= 5.0. We estimated the treatment effect within this population after conditioning on $E[Y|X, T = 0]$, where $E[Y|X, T = 0]$ was again estimated by fitting an outcome model within the untreated cohort and then used to predict disease risk for each individual within the full cohort. We then simulated four populations similar to the original population with the exception that we reduced the prevalence of treatment. We estimated the treatment effect within the original study cohort (simulated population with baseline treatment prevalence of 50%) after conditioning on DRSs that were estimated within the populations with reduced treatment prevalence.

Table 2 Simulation results for DRS estimation strategies to reduce bias amplification

DRS estimation method	Prevalence of treatment	Absolute bias
Same-sample estimation ^a	50%	0.44
Reduced tmt. prevalence ^b	10%	0.39
	5%	0.37
	1%	0.36
Historical population ^c	0%	0.35

Notes: ^aDRS estimated in the study cohort. ^bDRS estimated in a separate population with lower treatment prevalence. ^cDRS estimated in a population prior to treatment introduction.

Results in Table 2 illustrate that bias amplification can be reduced by estimating the DRS in a population where the prevalence of treatment is reduced compared to the original study cohort. Results further indicate that this reduction in bias is a function of the baseline prevalence of treatment in the population where the DRS is estimated. This simple example is not intended to be reflective of any practical settings, but is used to illustrate that conditioning on DRSs that are estimated in populations with reduced treatment prevalence only partially controls the induced confounding pathway (i.e. the induced correlation) that occurs between predictors of treatment and unmeasured confounders upon conditioning on treatment. This, in turn, reduces the overall confounding in the estimated treatment effect.

4.5 Challenges when comparing treatment to alternative therapies

In the previous discussion and simulated examples, we compared treated to untreated individuals to simplify the discussion. While comparing treated populations to no treatment or placebo is a question of interest in many studies, comparing treatments to alternative therapies (comparative effectiveness) is often more relevant for clinical decision making in pharmacoepidemiology. Comparing treatment to alternative therapies has many advantages with respect to avoiding the potential for confounding bias, but presents additional challenges when attempting to reduce bias amplification through the proposed DRS estimation strategies. This is because we can no longer avoid conditioning on treatment when estimating the DRS.

For example, if T_2 represents a comparator treatment and Y_2 the potential outcome under T_2 , then the DRS is defined as $E[Y_2|X]$ not $E[Y_0|X]$, where $E[Y_0|X]$ represents the baseline disease risk had individuals remained untreated (i.e. did not receive T_1 or T_2). Conditioning on $E[Y_0|X]$ induces prognostic balance for Y_0 by satisfying $Y_0 \perp X|E[Y_0|X]$, but does not necessarily satisfy $Y_2 \perp X|E[Y_0|X]$. Conditioning on $E[Y_0|X]$ when comparing alternative treatment therapies, therefore, requires the additional assumption that treatment effect heterogeneity for the comparator treatment is similar across populations that have the same distribution of baseline disease risk.

Because the DRS does not balance covariates across treatment groups, two groups that have the same distribution of baseline disease risk (i.e. $E[Y_0|X]$) do not necessarily imply that those treatment groups will have the same distribution of disease risk under the comparator treatment (i.e. $E[Y_2|X]$). Differences can arise because different covariate distributions across the two treatment groups can potentially result in differences in treatment effect heterogeneity for the comparator treatment across the two populations. If there is no treatment effect heterogeneity, or it is reasonable to assume that treatment effect heterogeneity for the comparator treatment is similar across populations with the same distribution of baseline disease risk, then treatment groups with the same distribution of baseline disease risk will also have the same distribution of disease risk under the comparator treatment. Under this assumption, the proposed estimation strategies for baseline disease risk (i.e. $E[Y_0|X]$) can potentially reduce bias amplification when

comparing alternative treatment therapies while at the same time control for the confounding effects of measured confounders.

5 Empirical example: COX-2 inhibitors vs NSAIDs

We compared new-users of the COX-2 inhibitor celecoxib to new-users of non-selective NSAIDs (ibuprofen or diclofenac) in reducing gastrointestinal (GI) complications over a 60-day follow-up period. New-users were defined as individuals who initiated a celecoxib or NSAID between the years 2006 and 2008 after having no prescription of these medications during a 6-month washout period. GI complication was defined as a hospitalization for a GI hemorrhage or peptic ulcer disease or an outpatient visit for a GI hemorrhage. Analyses were performed in the MarketScan Commercial Claims and Encounters and Medicare Supplementary and Coordination of Benefit database (Truven Healthcare, Inc.), a large US employer-based insurance claims database. The database contains patient billing information for in- and outpatient procedures and diagnoses, pharmacy medication dispensing, and enrollment information for enrolled employees, spouses, dependents, and retirees. We included all individuals who were continuously enrolled in MarketScan for at least 12 months prior to drug initiation. All demographic and clinical covariates were defined during the 6 months prior to drug initiation and are described in Table 3.

We chose this example for a few reasons. First, observational studies comparing COX-2 inhibitors to NSAIDs using claims data are known to contain some degree of unmeasured confounding due to a lack of information on important confounding factors (e.g. smoking status, body mass index, alcohol consumption, etc.) [16, 25]. Further, previous studies have shown strong confounding by indication when comparing COX-2 inhibitors to NSAIDs in preventing GI complication [26, 27]. The likelihood for confounding by indication in this setting is supported by strong differences in the distribution of baseline covariates across treatment groups shown in Table 3. Finally, we can use results from clinical trials as a benchmark when evaluating the performance of various methods for DRS estimation [28, 29].

The effect of initiating celecoxib vs NSAIDs was estimated using a Cox proportional hazards model after stratifying on the estimated DRS. We also estimated the treatment effect after stratifying on the estimated PS for comparison. We estimated the DRS in two different populations. We first estimated the DRS (DRS1) within the comparator group of NSAID new-users (ibuprofen or diclofenac). We also estimated the DRS

Table 3 Covariate summary

Baseline covariates	celecoxib (<i>N</i> = 4,606)	NSAID (<i>N</i> = 17,175)
Age (mean)	60.7	55.2
Sex (female)	61.2%	60.2%
Rheumatoid arthritis	2.9%	1.2%
Osteoarthritis	27.8%	12.9%
Backache	24.1%	17.0%
Chest pain	15.0%	12.2%
Osteoporosis	4.0%	2.1%
Atrial fibrillation	9.1%	5.7%
Abdominal pain	21.5%	17.1%
Peptic ulcer, GERD, or esophageal ulcer	10.7%	7.3%
Heart disease	20.6%	15.5%
Hypertension	40.3%	32.8%
PPI	14.8%	9.7%
Warfarin use	6.5%	3.4%
Pain	32.5%	22.3%
Six or more prescription drugs in prior year	69.8%	57.9%

(DRS2) in a population of untreated individuals (celecoxib and NSAID non-users). Because surveillance between non-initiators and new-users can differ, we identified a group of healthcare-seeking individuals who had an outpatient physicians visit after a 6-month washout period where no celecoxib or NSAID was used. We required that individuals have an outpatient physicians visit in order to minimize potential differences in covariate surveillance between non-initiators and new-users. For each of the DRS models, we calculated the c-statistic and Hosmer–Lemeshow test statistic within a validation set of controls (i.e. a subset of NSAID new-users) to evaluate the discrimination and calibration of the fitted DRS model. Individuals were censored if they lost any part of coverage during follow-up.

PS stratification and both estimation methods for the DRS gave similar results in terms of the estimated hazard ratio (Table 4). Although both PS and DRS stratification moved the effect estimate in the expected direction, none of these methods resulted in effect estimates that are consistent with clinical trials which have shown COX-2 inhibitors to have a significant protective effect against GI complications compared to non-selective NSAIDs [28, 29]. These discrepancies indicate the existence of strong unmeasured confounding. In this example, we did not see an improvement in a reduction of bias amplification when estimating the DRS within an untreated population where the prevalence of both treatments (celecoxib and NSAIDs) were reduced. Estimating the DRS within NSAID new-users (DRS1) resulted in better model fit with improved discrimination (c-statistic) and calibration (Hosmer–Lemeshow test) compared to estimating the DRS within an untreated population (DRS2).

Table 4 Results comparing celecoxib vs NSAIDs

	HR	95% CI	C-statistic	Hosmer–Lemeshow ^a
Unadjusted	1.21	(0.93, 1.50)	–	–
PS stratification	0.98	(0.73, 1.31)	–	–
DRS1 stratification ^b	0.97	(0.72, 1.30)	0.65	0.21
DRS2 stratification ^c	1.02	(0.75, 1.37)	0.61	0.01

Notes: ^a*p*-value for the Hosmer–Lemeshow test for goodness of fit; ^bdisease risk score estimated in a comparator group of NSAID new-users; ^cdisease risk score estimated in a population of untreated individuals (celecoxib and NSAID non-users).

6 Conclusions

We have given a concise overview of bias amplification and have discussed how bias amplification results from controlling induced correlations that occur between predictors of treatment and unmeasured confounders when conditioning on treatment. We have further discussed ways in which bias amplification can potentially be reduced using alternative estimation strategies for the DRS. Estimating the DRS in an outside population using either historical or contemporary data allows researchers to potentially reduce bias amplification caused by both instrumental variables and measured confounders. In theory, alternative estimation strategies for the DRS provide a way to avoid conditioning on measured instrumental variables and may reduce the cost of controlling for measured confounders that have strong effects on treatment and weak effects on the outcome.

Although estimating the DRS in alternative populations has the potential to reduce bias amplification, applying these estimation strategies within administrative claims databases has limitations as illustrated in the empirical example comparing new-users of the COX-2 inhibitor celecoxib to new-users of non-selective NSAIDs. As discussed previously, estimating the DRS in untreated individuals outside of the defined study cohort has the potential for covariate and outcome misclassification due to differential surveillance of individuals within the database. Further, estimating the DRS within populations outside the study cohort requires the assumption that the effects of baseline covariates on the outcome are not modified in the population where disease risk is estimated. Estimation of disease risk in a lower risk population

(non-initiators) can lead to problems with extrapolation of the DRS and possibly non-linear associations with risk. Additional research is needed to evaluate the performance of DRSs when using external data to estimate the DRS in settings specific to pharmacoepidemiology and large database research.

Acknowledgments: This work was supported by a grant (R01 AG023178) from the National Institute on Aging. The authors also thank the Editor and reviewers for their helpful comments.

Appendix

Consider the causal scenario described in Figure 6(a) where $\lambda_5 = \frac{\partial}{\partial x} E[U|x, t]$ is the regression coefficient of X in a regression of U on X and T . Let λ_5^* represent the same regression coefficient, but within a population with reduced treatment prevalence compared to the original study population (Figure 6(b)).

$$\begin{aligned} \lambda_5 &= \rho_{UX|T} \frac{\sigma_{U|T}}{\sigma_{X|T}} \\ &= \left(\frac{\rho_{UX} - \rho_{UT}\rho_{XT}}{\sqrt{1 - \rho_{UT}^2}\sqrt{1 - \rho_{XT}^2}} \times \frac{\sigma_U \sqrt{1 - \rho_{UT}^2}}{\sigma_X \sqrt{1 - \rho_{XT}^2}} \right) \\ &= \left(\frac{-\rho_{UT}\rho_{XT}}{\sqrt{1 - \rho_{UT}^2}\sqrt{1 - \rho_{XT}^2}} \times \frac{\sigma_U \sqrt{1 - \rho_{UT}^2}}{\sigma_X \sqrt{1 - \rho_{XT}^2}} \right) \end{aligned}$$

In the above equations, ρ_{UX} and $\rho_{UX|T}$ represent the correlation between U and X and the correlation between U and X after conditioning on T , respectively. The term $\sigma_{U|T}$ is the standard deviation of U conditional on T . Other terms can be interpreted similarly. Since $\rho_{XT} = \frac{\text{cov}(X,T)}{\sigma_X \sigma_T}$ approaches 0 as the prevalence of treatment approaches 0, it follows that $|\lambda_5^*| \leq |\lambda_5|$.

References

1. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41–55.
2. Stürmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariate methods. *J Clin Epidemiol* 2006;59:437–47.
3. Hansen BB. The prognostic analogue to the propensity score. *Biometrika* 2008;95:481–8.
4. Arbogast PG, Ray WA. Use of disease risk scores in pharmacoepidemiologic studies. *Stat Methods Med Res* 2009;18:67–80.
5. Arbogast PG, Ray WA. Performance of disease risk scores, propensity scores, and traditional multivariable outcome regression in the presence of multiple confounders. *Am J Epidemiol* 2011;174:613–20.
6. Cadarette SM, Gagne JJ, Solomon DH, Katz JN, Stürmer T. Confounder summary scores when comparing the effects of multiple drug exposures. *Pharmacoepidemiol Drug Saf* 2010;19:2–9.
7. Glynn RJ, Gagne JJ, Schneeweiss S. Role of disease risk scores in comparative effectiveness research with emerging therapies. *Pharmacoepidemiol Drug Saf* 2012;21:138–47.
8. Tadrous M, Gagne JJ, Stürmer T, Cadarette SM. Disease risk score as a confounder summary method: systematic review and recommendations. *Pharmacoepidemiol Drug Saf* 2013;22:122–9.
9. Leacy FP, Stuart EA. On the joint use of propensity and prognostic scores in estimation of the average treatment effect on the treated: a simulation study. *Stat Med* 2013 [Epub ahead of print].
10. Bhattacharya J, Vogt WB. Do instrumental variables belong in propensity scores? NBER technical working paper no. 343. National Bureau of Economic Research, Cambridge, MA, 2007.
11. Wooldridge J. Should instrumental variables be used as matching variables? Technical report. East Lansing, MI: Michigan State University, 2009.

12. Pearl J. On a class of bias-amplifying variables that endanger effect estimates. In: Proceedings of the twenty-sixth conference on uncertainty in artificial intelligence, AUAI, Corvallis, OR, 2010:417–24.
13. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. *Am J Epidemiol* 2006;163:1149–56.
14. Myers JA, Rassen JA, Gagne JJ, Huybrechts KF, Schneeweiss S, Rothman KJ, et al. Effects of adjusting for instrumental variables on bias and precision of effect estimates. *Am J Epidemiol* 2011;174:1213–22.
15. Pearl J. Invited commentary: understanding bias amplification. *Am J Epidemiol* 2011;174:1223–7;discussion 1228–9.
16. Brookhart MA, Stürmer S, Glynn RJ, Rassen J, Schneeweiss S. Confounding control in healthcare database research: challenges and potential approaches. *Med Care* 2010;48:S114–20.
17. Wright S. Correlation and causation. *J Agric Res* 1921;20:557–85.
18. Lunt M, Avorn J, Glynn RJ, Rothman KJ, Stürmer T. Propensity score calibration in the absence of surrogacy. *AJE* 2012;175:1294–302.
19. Pearl J. Linear models: a useful microscope for causal analysis. *J Causal Inference* 2013;1:155–70.
20. Rubin D. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol* 1974;66:688–701.
21. Little R, Rubin D. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annu Rev Public Health* 2000;21:121–45.
22. Pearl J. *Causality: models, reasoning, and inference*, 2nd ed. New York, NY: Cambridge University Press, 2009.
23. Richardson TS, Robins JM. Single world intervention graphs (SWIGS): a unification of the counterfactual and graphical approaches to causality. Working paper number 128. Center for Statistics and the Social Sciences, 2013.
24. Miettinen OS. Stratification by a multivariate confounder score. *Am J Epidemiol* 1976;104:609–20.
25. Ray WA. Evaluating medication effects outside clinical trials: new-user designs. *Am J Epidemiol* 2003;158:915–20.
26. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology* 2006;17:268–75.
27. Patino FG, Allison J, Olivieri J, Mudano A, Juarez L, Person S, et al. The effects of physician specialty and patient comorbidities on the use and discontinuation of coxibs. *Arthritis Rheum* 2003;49:293–9.
28. Silverstein FE, Faich G, Goldstein JL, Simon LS, Pincus T, Whelton A, et al. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: a randomized controlled trial. Celecoxib long-term arthritis safety study. *J Am Med Assoc* 2000;284:1247–55.
29. Bombardier C, Laine L, Reicin A, Shapiro D, Burgos-Vargas R, Davis B, et al. Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. VIGOR study group. *N Engl J Med* 2000;343:1520–8.