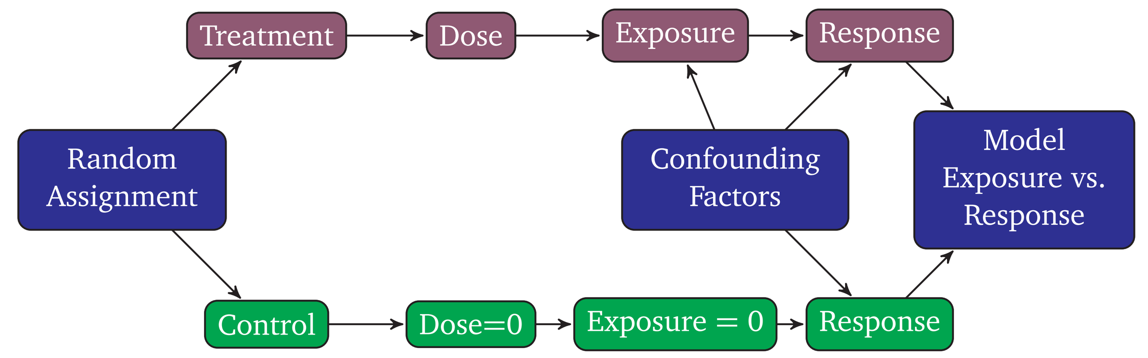
Evaluating effectiveness of case-matching for exposure-response analysis

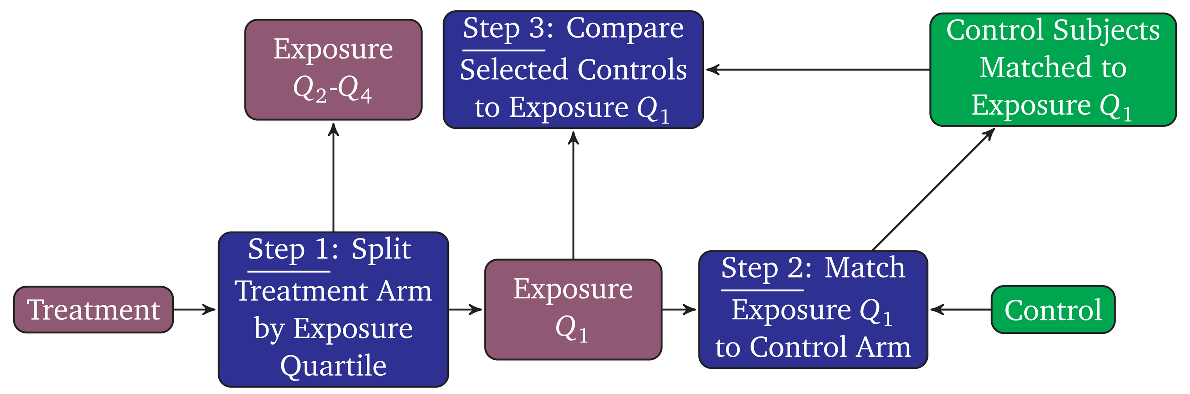
Jayson D. Wilbur[1], Manish Gupta[2], Chaitali Passey[2] and Amit Roy[2]
[1]*Metrum Research Group LLC, Tariffville CT, USA* [2]*Bristol-Myers-Squibb, Lawrenceville NJ, USA*

Case-matching in exposure-response analysis

Characterization of exposure-response relationships can be challenging in the presence of confounding factors that affect both pharmacokinetic properties as well as the response.


Treatment
Dose
Exposure
Response
Random Assignment
Confounding Factors
Model Exposure vs. Response
Control
Dose=0
Exposure = 0
Response

In such situations, case-matching has been proposed to select control arm subjects for inclusion in the analysis [1]. In particular, sponsors have been asked to use the relationship between measured covariates and exposure, as observed in the treatment arm, to match subjects in the lowest exposure quartile to subjects in the control arm.


Exposure Q_2-Q_4
Step 3: Compare Selected Controls to Exposure Q_1
Control Subjects Matched to Exposure Q_1
Treatment
Step 1: Split Treatment Arm by Exposure Quartile
Exposure Q_1
Step 2: Match Exposure Q_1 to Control Arm
Control

Should we expect matched controls to have low exposure if treated?

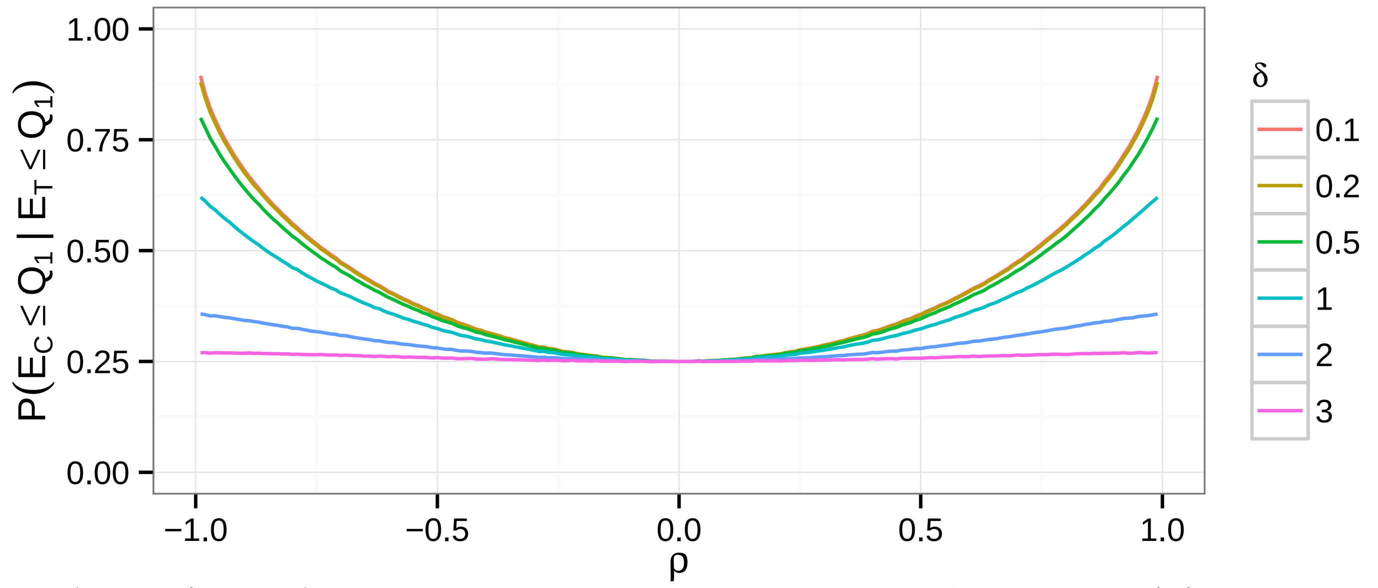
Suppose that for the population of interest, the joint distribution between predictor Z (e.g., a propensity score) and exposure when treated E, is described by a standard bivariate normal distribution with correlation ρ; and consider a treated subject T with covariate-exposure pair (Z_T, E_T).

We are interested in the probability that an independent subject C in the control arm with a similar predictor value $Z_C \in \{z : |z - Z_T| < \delta\sigma_Z\}$ would have experienced similar exposure, had the subject been treated.
(Note that if Z does represent the propensity score, then δ would typically be referred to as the caliper width [2].)

In particular, we are interested in $P(E_C \le Q_1 | E_T \le Q_1)$, where Q_1 represents the first quartile of E.

$$
\begin{aligned}
F(e_C, z_C, z_T, e_T) &= F(e_C|z_C, z_T, e_T)F(z_C|z_T, e_T)F(z_T|e_T)F(e_T) \\
&= F(e_C|z_C)F(z_C|z_T)F(z_T|e_T)F(e_T) \\
&= \Phi\left(\frac{e_C - \rho z_C}{\sqrt{1-\rho^2}}\right)\frac{\min\left(1, [\Phi(z_C) - \Phi(z_T - \delta)]^+\right)}{\Phi(z_T + \delta) - \Phi(z_T - \delta)}\Phi\left(\frac{z_T - \rho e_T}{\sqrt{1-\rho^2}}\right)\Phi(e_T) \\
P(E_C \le Q_1 | E_T \le Q_1) &= \lim_{e_C, e_T \to Q_1 \;\; z_C, z_T \to \infty} \frac{F(e_C, z_C, z_T, e_T)}{F(e_T)} \\
&= \lim_{e_C, e_T \to Q_1 \;\; z_C, z_T \to \infty} \Phi\left(\frac{e_C - \rho z_C}{\sqrt{1-\rho^2}}\right)\frac{\min\left(1, [\Phi(z_C) - \Phi(z_T - \delta)]^+\right)}{\Phi(z_T + \delta) - \Phi(z_T - \delta)}\Phi\left(\frac{z_T - \rho e_T}{\sqrt{1-\rho^2}}\right) \\
&= \lim_{z \to \infty} \Phi\left(\frac{Q_1 - \rho(z + \gamma)}{\sqrt{1-\rho^2}}\right)\frac{\Phi(z + \gamma) - \Phi(z - \delta)}{\Phi(z + \delta) - \Phi(z - \delta)}\Phi\left(\frac{z - \rho Q_1}{\sqrt{1-\rho^2}}\right)
\end{aligned}
$$

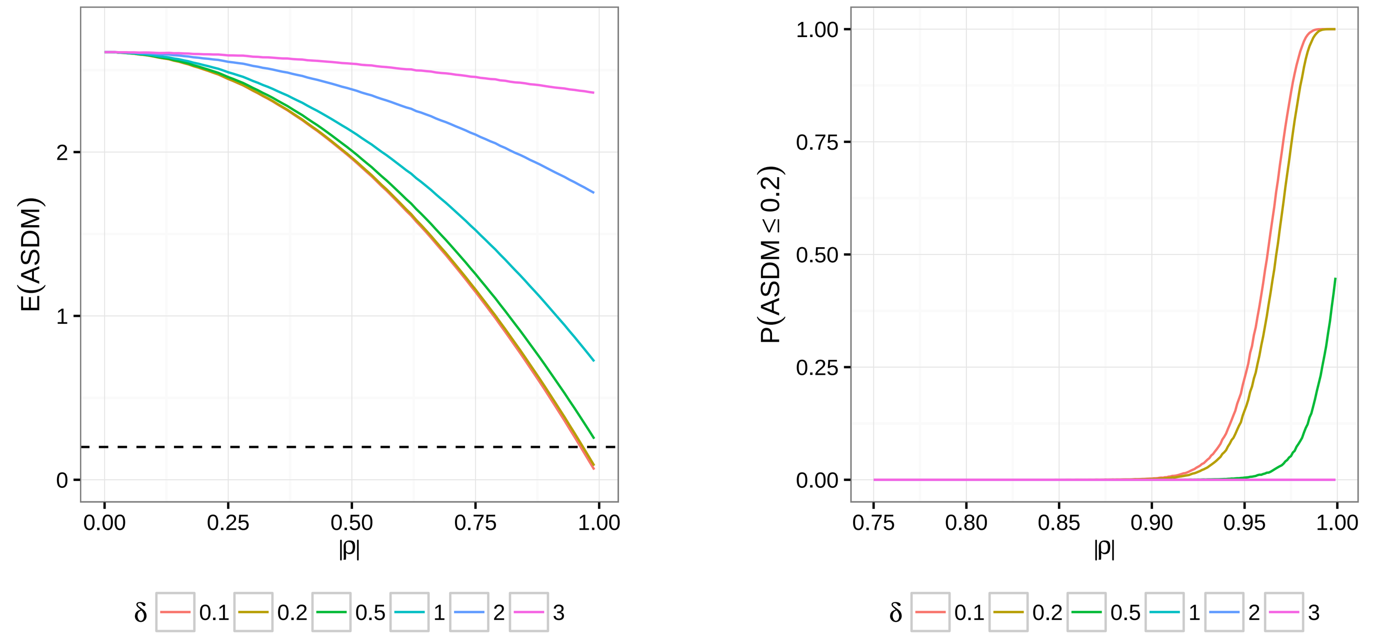
where $\gamma = (z_C - z_T) \in (-\delta, +\delta)$. For the extreme cases of $\rho = 0$ and $|\rho| = 1$, it can be seen that $P(E_C \le Q_1 | E_T \le Q_1) = \Phi(Q_1) = 0.25$ and 1, respectively. Further, we show by simulation that $P(E_C \le Q_1 | E_T \le Q_1)$ is an increasing function of $|\rho|$ and a decreasing function of δ.


P(E_C ≤ Q_1 | E_T ≤ Q_1)
ρ
δ
0.1
0.2
0.5
1
2
3

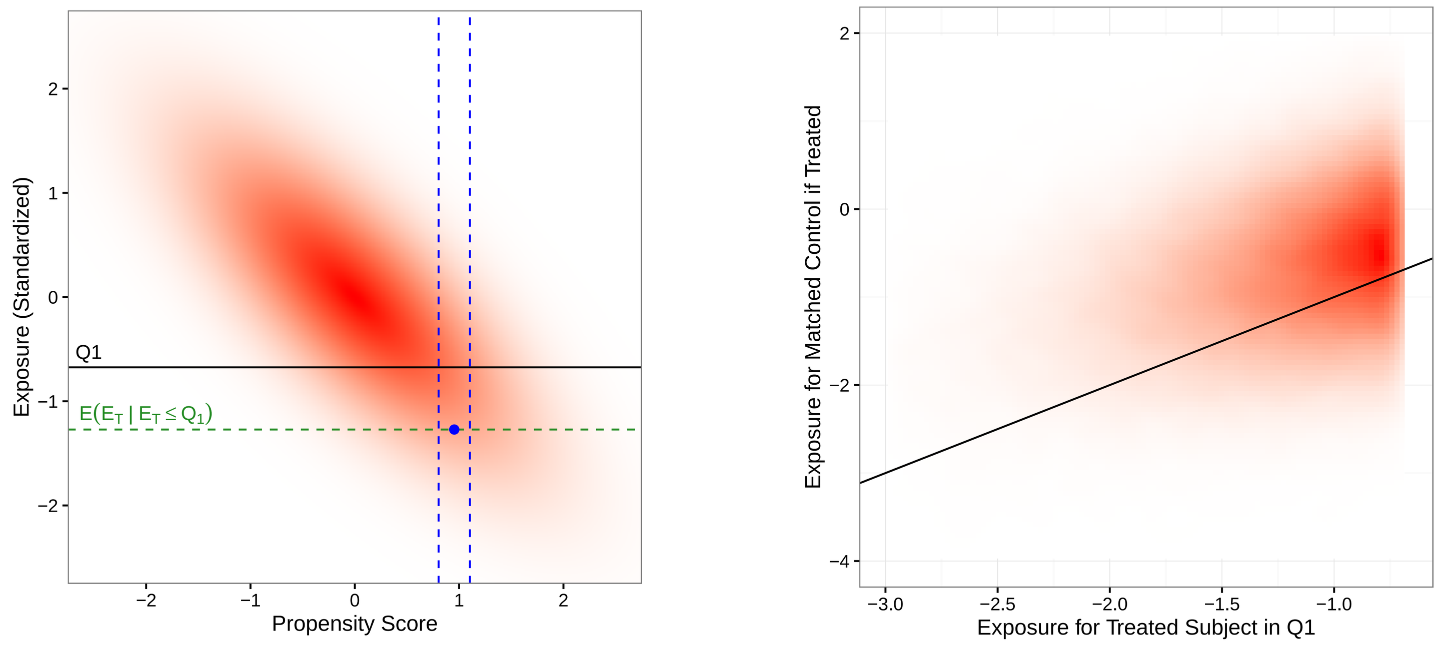
While $P(E_C \le Q_1 | E_T \le Q_1)$ is lower than one might expect (except possibly for $\delta \le 0.2$ and $|\rho| \ge 0.9$), the effectiveness of case-matching is not typically judged in terms of the similarity of individual matches, but by the similarity of the groups in aggregate. In particular, case-matching is typically deemed effective if the absolute standardize difference in means is less than 0.2 for a collection of observed covariates (i.e., ASDM=$|\bar{x}_{Q_1} - \bar{x}_{Matched}|/s_{Q_1} \le 0.2$).

Balance of the exposure distributions after matching

We obtain estimates for $E(ASDM)$ and $P(ASDM \le 0.2)$ by simulation (using $n = 125$ for Q_1 of the treatment arm) and see that the distribution of exposure remains unbalanced, except possibly for $\delta \le 0.2$ and $|\rho| \ge 0.9$.


E(ASDM)
|ρ|
P(ASDM ≤ 0.2)
|ρ|
δ
0.1
0.2
0.5
1
2
3

To understand the source of this imbalance, consider the density plot in the left panel below, illustrating a negative correlation between exposure and the propensity score for assignment to Q_1 of exposure in the treatment arm.


Exposure (Standardized)
Q1
E(E_T | E_T ≤ Q_1)
Propensity Score
Exposure for Matched Control if Treated
Exposure for Treated Subject in Q1

A subject in Q_1, such as the example indicated by the blue dot, is matched to a subject in the control arm with a propensity score within the caliper indicated by the blue vertical dashed lines.

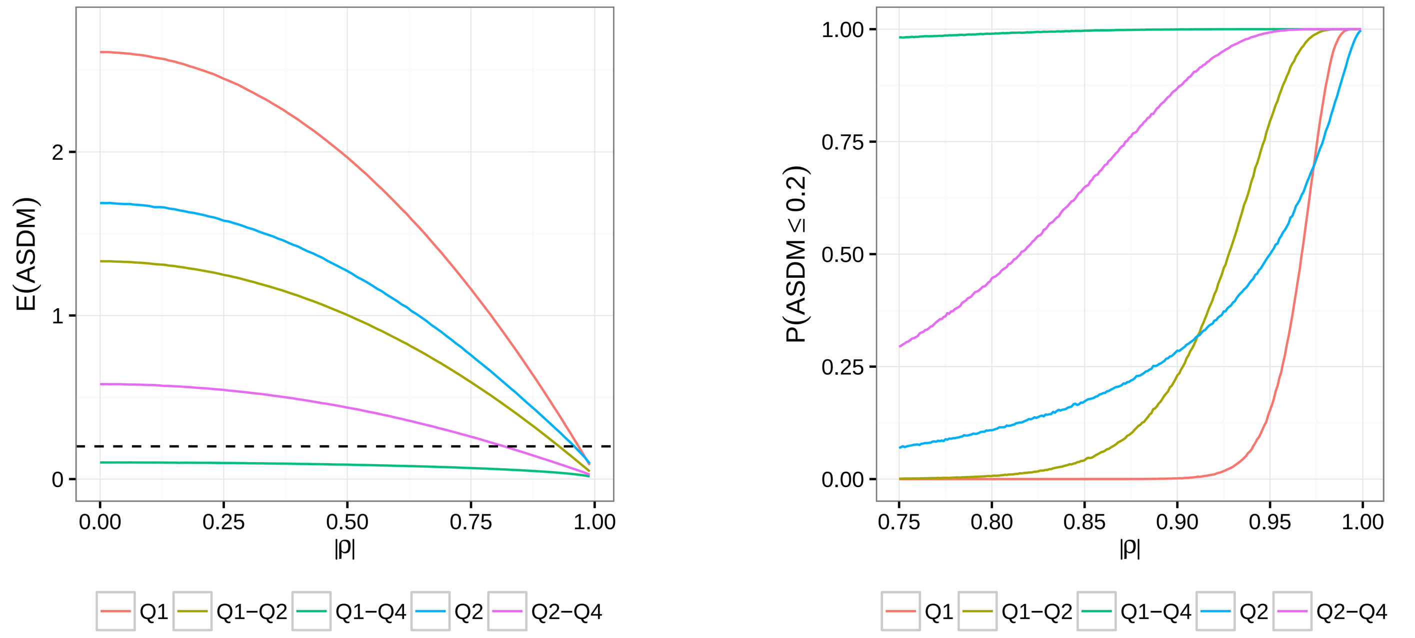
It can be seen that the exposure distribution for subjects within the caliper extends beyond Q_1, and that if treated, the matched control would likely have higher exposure than its treated counterpart.

Further, subjects in Q_1 are in the upper tail of the propensity score distribution. Thus, the distribution of propensity scores within the caliper is not uniform, with bias toward lower propensity scores, compounding the problem.

These observations are borne out by the plot in the right panel above which shows that matched controls would tend to have slightly higher exposure than their counterparts in Q_1 of the treatment arm (given the assumption of a bivariate normal distribution with negative correlation between exposure and the propensity score).

It is further noted that as the sample size increases, $P(ASDM \le 0.2)$ will actually decrease as the estimate for the mean exposure of the matched controls (if treated) converges to its expectation, which is necessarily higher than $E(E_T | E_T \le Q_1)$, due to the small bias expected from each match.

To further illustrate that the stratification of exposure is the reason for the bias, we fix $n = 125$ and $\delta = 0.2$, and obtain estimates of $E(ASDM)$ and $P(ASDM \le 0.2)$ for different subsets of the treatment arm and note the effect.


E(ASDM)
|ρ|
P(ASDM ≤ 0.2)
|ρ|
Q1
Q1−Q2
Q1−Q4
Q2
Q2−Q4

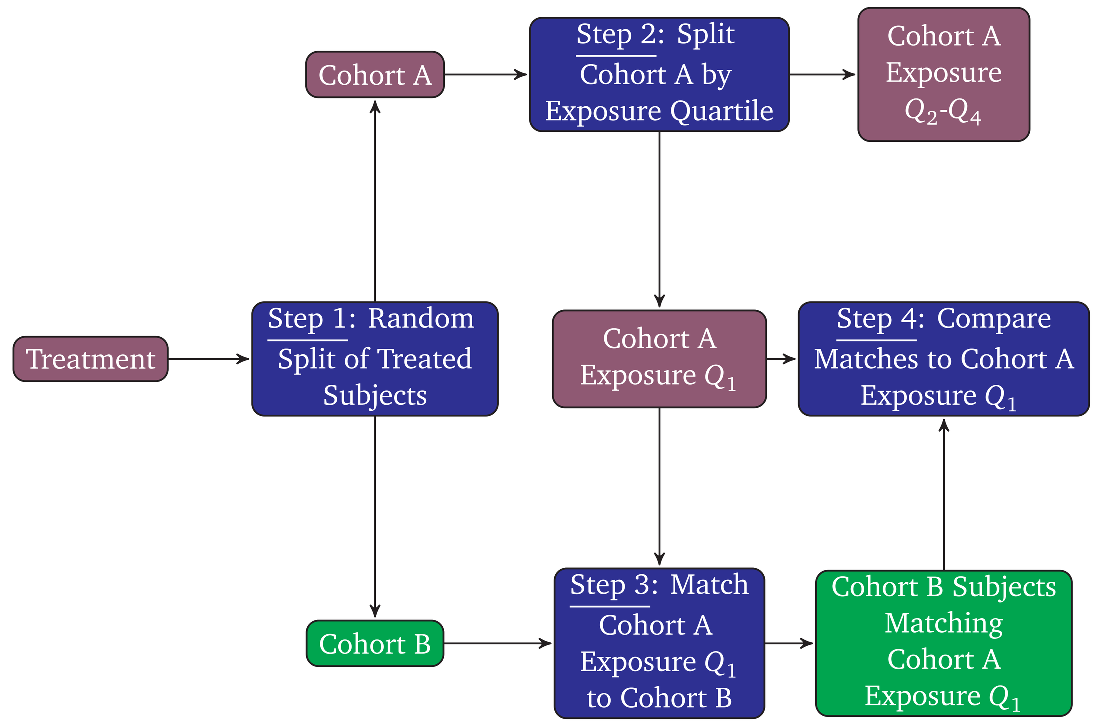
Methods for evaluating the effectiveness of case-matching

While case-matching in exposure-response analysis does not ensure balance with respect to exposure, it does identify a subset of the control arm more similar to the subjects in the treatment arm in the lowest exposure quartile.

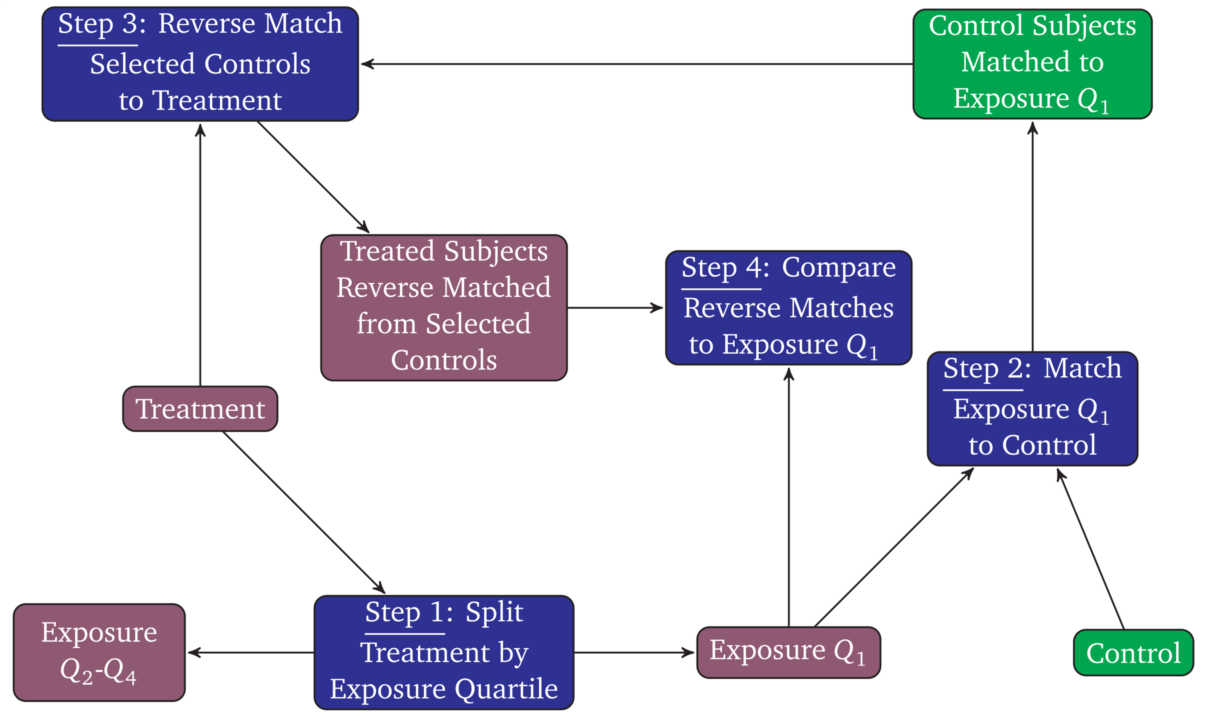
Because exposure is unobservable in the control arm, and may have a non-linear relationship with the propensity score, the correlation between the propensity score and exposure in the treatment arm may be insufficient for evaluating the effectiveness of case-matching procedures for particular studies.

Therefore, we propose the following two methods for evaluating the effectiveness of case-matching in exposure-response analyses [3].

Method #1: Matching within Treatment Arm


Cohort A
Step 2: Split Cohort A by Exposure Quartile
Cohort A Exposure Q_2-Q_4
Treatment
Step 1: Random Split of Treated Subjects
Cohort A Exposure Q_1
Step 4: Compare Matches to Cohort A Exposure Q_1
Cohort B
Step 3: Match Cohort A Exposure Q_1 to Cohort B
Cohort B Subjects Matching Cohort A Exposure Q_1

Method #2: Reverse Matching Selected Controls back to the Treatment Arm


Step 3: Reverse Match Selected Controls to Treatment
Control Subjects Matched to Exposure Q_1
Treated Subjects Reverse Matched from Selected Controls
Step 4: Compare Reverse Matches to Exposure Q_1
Treatment
Step 2: Match Exposure Q_1 to Control
Exposure Q_2-Q_4
Step 1: Split Treatment by Exposure Quartile
Exposure Q_1
Control

Simulation study for method validation

The validity of these two methods was assessed for several simulated scenarios using various sample sizes, numbers of continuous covariates and magnitude of correlation between the covariates themselves and with exposure.

For each subject in the treatment and control arms we generate the $(p+1) \times 1$ vector

$$\boldsymbol{X}_i = \left[\; X_{i0} \mid X_{i1} \;\; \cdots \;\; X_{ip} \;\right]$$

consisting of a measure of exposure, X_{i0} and the p covariates, $X_{i1}, \ldots, X_{ip}$, according to

$$N_{p+1}\left(\boldsymbol{\mu} = \boldsymbol{0}_{p+1}, \boldsymbol{\Sigma} = \sigma^2\left[(1-\rho)\boldsymbol{I}_{p+1} + \rho\boldsymbol{1}_{p+1}\boldsymbol{1}'_{p+1}\right]\right)$$

fixing $\sigma^2 = 0.5$ and varying subjects per arm $N = 100, 200, 500$, covariates $p = 5, 10, 20$ and $\rho = 0.25 \ldots 0.99$.

Treated subjects in the lowest quartile of the exposure distribution were matched to subjects in the control arm using a logistic propensity score model [4] that included all covariates. Matches were selected at random from candidates with propensity scores within a caliper of $\delta = 0.2$ times the standard deviation of the propensity score distribution.

References

[1] Yang, J., Zhao, H., Garnett, C., Rahman, A., Gobburu, J.V., Pierce, W., Schechter, G., Summers, J., Keegan, P., Booth, B. and Wang, Y. The combination of exposure-response and case-control analyses in regulatory decision making. *J Clin Pharmacol* **53** (2013):160–6.

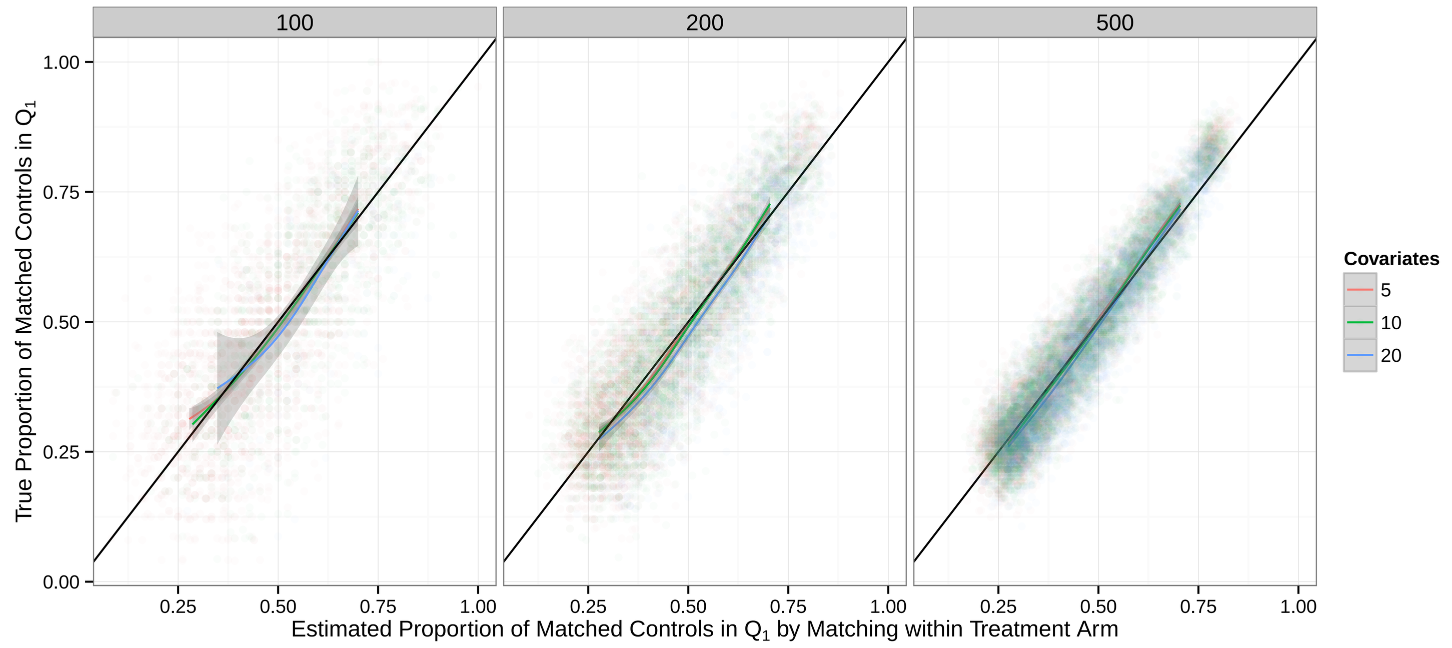
[2] Austin, P.C. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical statistics* **10** (2011):150–161.

[3] Wilbur, J.D., Gupta, M., Passey, C. and Roy, A. Evaluating effectiveness of case-matching for exposure-response analysis (2016). *PAGE.* *URL* `http://page-meeting.org/pdf_assets/9662-Wilbur-EvalCaseControl-PAGE16.pdf`

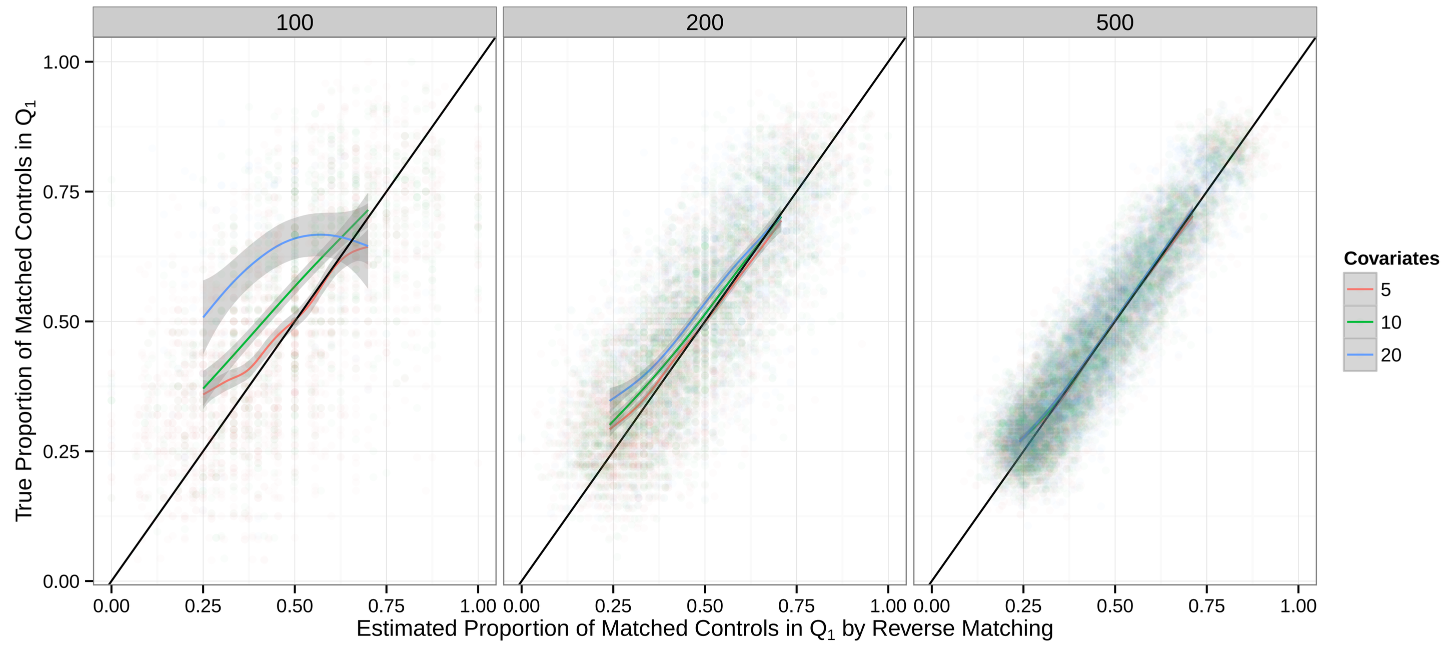
[4] *D'Agostino, Jr, R.B. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group.* Stat Med *17 (1998):2265–81.*

Results

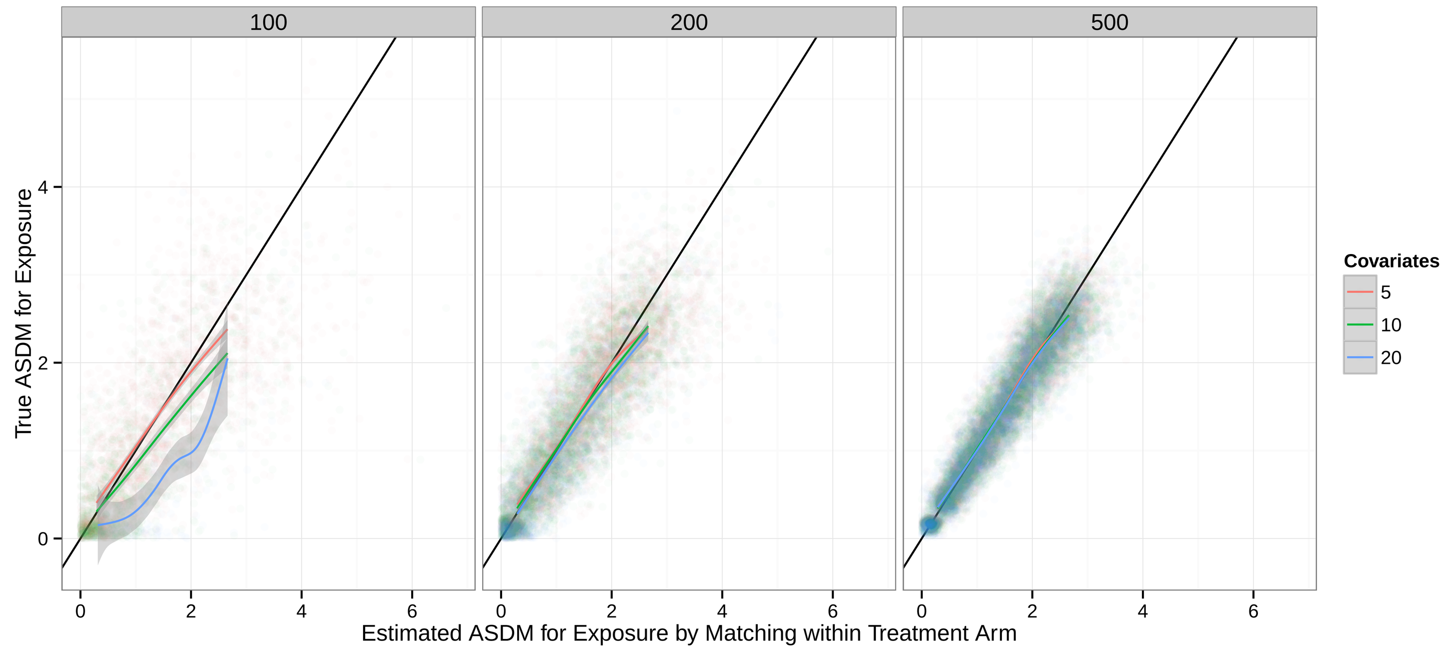
Reverse matching provides unbiased estimates for the proportion of matched controls with low exposure.


100
200
500
True Proportion of Matched Controls in Q_1
Estimated Proportion of Matched Controls in Q_1 by Matching within Treatment Arm
Covariates
5
10
20

Matching within the treatment arm provides unbiased estimates for the proportion of matched controls with low exposure when $n \ge 200$ per arm. Note that estimates based on matching within a treatment arm of $n = 100$ are the result of matching the very few (i.e., 12-13) subjects in the lowest exposure quartile of the split treatment arm.


100
200
500
True Proportion of Matched Controls in Q_1
Estimated Proportion of Matched Controls in Q_1 by Reverse Matching
Covariates
5
10
20

Matching within the treatment arm provides unbiased estimates of the ASDM for exposure when $n \ge 200$ per arm.


100
200
500
True ASDM for Exposure
Estimated ASDM for Exposure by Matching within Treatment Arm
Covariates
5
10
20

Conclusions

In the context of exposure-response analysis, case-matching is increasingly being used to identify a subset of the control arm similar those in the treatment arm with low exposure. However, case-matching does not ensure that the matched controls would have similarly low exposure if treated. Therefore, even after adjustment by case-matching, estimates for the steepness of the exposure-response relationship may remain biased.

The stratification of the exposure distribution prior to matching was shown to be the reason for the residual imbalance in exposure. Therefore, these results do not imply an expectation of imbalance for unobserved covariates.

Two methods for evaluating the effectiveness of case-matching in the context of exposure-response analysis are proposed and it is demonstrated that both methods would be useful as part of a case-matching evaluation strategy.

It is recommended that the effectiveness of case-matching be evaluated prior to performing exposure-response analysis, to ensure that the matching can be expected to result in balanced distributions for exposure.