Combining distributed regression and propensity scores: a doubly privacy-protecting analytic method for multicenter research

Senghee Toh1
Robert Wellman2
Robert Yates Coley2
Casie Horgan1
Jessica Sturtevant1
Erick Moyneur3
Cheri Janning4
Roy Pardee2
Karen J Coleman5
David Arterburn2
Kathleen McTigue6
Jane Anau2
Andrea J Cook2

On behalf of the PCORnet Bariatric Study Collaborative

1Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA; 2Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA; 3StatLog Econometrics, Inc., Montreal, QC, Canada; 4Duke Clinical and Translational Science Institute, Durham, NC, USA; 5Kaiser Permanente Southern California, Pasadena, CA, USA; 6Department of Medicine, University of Pittsburgh, Pittsburgh, PA, USA

This article was published in the following Dove Press journal: Clinical Epidemiology

Purpose: Sharing of detailed individual-level data continues to pose challenges in multicenter studies. This issue can be addressed in part by using analytic methods that require only summary-level information to perform the desired multivariable-adjusted analysis. We examined the feasibility and empirical validity of 1) conducting multivariable-adjusted distributed linear regression and 2) combining distributed linear regression with propensity scores, in a large distributed data network.

Patients and methods: We compared percent total weight loss 1-year postsurgery between Roux-en-Y gastric bypass and sleeve gastrectomy procedure among 43,110 patients from 36 health systems in the National Patient-Centered Clinical Research Network. We adjusted for baseline demographic and clinical variables as individual covariates, deciles of propensity scores, or both, in three separate outcome regression models. We used distributed linear regression, a method that requires only summary-level information (specifically, sums of squares and cross products matrix) from sites, to fit the three ordinary least squares linear regression models. A comparison set of analyses that used pooled deidentified individual-level data from sites served as the reference.

Results: Distributed linear regression produced results identical to those from the corresponding pooled individual-level data analysis for all variables in all three models. The maximum numerical difference in the parameter estimate or standard error for all the variables was $3 \times 10^{-11}$ across three models.

Conclusion: Distributed linear regression analysis is a feasible and valid analytic method in multicenter studies for one-time continuous outcomes. Combining distributed regression with propensity scores via modeling offers more privacy protection and analytic flexibility.

Keywords: distributed regression, propensity score, distributed data networks, privacy-protecting methods

Introduction

It is increasingly common and necessary to conduct multicenter, population-based studies with data from a large number of geographically and demographically diverse individuals.1–6 However, it is not always possible or desirable to share deidentified individual-level data sets in multicenter studies. For example, some data partners may be concerned about potential breaches of patient privacy, while others may be worried about unauthorized uses of their data, incorrect analysis or interpretation of the data, or inadvertent disclosures of sensitive corporate or institutional information.7–10 Certain health plans or delivery systems, which collect electronic health data as part of clinical care, have contractual agreements with their patients that restrict sharing
of individual-level data for purposes other than patient care. Even when patients, health plans, and delivery systems are amenable to sharing individual-level data, sometimes the time and resources required to formalize data-sharing agreements can significantly delay a research project or make it infeasible.

There have been ongoing efforts to develop and apply more privacy-protecting analytic and data-sharing methods that do not require transferring of individual-level data sets in multicenter studies. For example, methods that leverage summary scores (eg, propensity scores and disease risk scores) are appealing because they allow researchers to adjust for a large number of covariates without having to share highly granular information. Prior studies have demonstrated the feasibility and validity of using only summary-level information from participating sites to perform matched or stratified analysis with summary scores for fixed exposures and one-time binary and time-to-event outcomes. Meta-analysis of site-specific effect estimates has also been shown to be a viable analytic option in certain multicenter studies. Distributed regression, which has theoretical advantages of being computationally equivalent to pooled individual-level data regression, has also been used in some multicenter studies.

There is a continued need for more methodological development and empirical assessment of these more privacy-protecting analytic methods in real-world multicenter studies. We performed a study to empirically assess the feasibility and validity of 1) conducting distributed linear regression analysis in a large national distributed data network; and 2) combining distributed regression with propensity scores, a data dimension reduction technique, to achieve more privacy protection and analytic flexibility.

Materials and methods

Setting

The National Patient-Centered Clinical Research Network (PCORnet) is a distributed data network designed to support studies that address questions important to patients and other stakeholders. It includes 13 Clinical Data Research Networks (CDRNs), 20 People-Powered Research Networks, two Health Plan Research Networks, and a coordinating center. The CDRNs, each of which includes multiple participating health systems, provide access to electronic health record (EHR) or administrative claims data from more than 100 million individuals and over 40 million patients who could be recruited into pragmatic clinical trials. PCORnet employs a common data model to standardize the data extracted from the EHRs or administrative claims of the participating health systems. The data domains extracted from the EHRs included patient demographics, diagnoses and procedures, vital signs, and laboratory test results.

Study cohort

The PCORnet Bariatric Study was initiated in 2016 to evaluate the comparative effectiveness and safety of three commonly performed bariatric procedures – adjustable gastric band, Roux-en-Y gastric bypass, and sleeve gastrectomy. The study identified patients who underwent a primary bariatric procedure in participating health systems from January 1, 2005, to September 30, 2015. Patients must meet the following criteria to be eligible for the study: 1) aged 21–79 years at the index procedure; 2) body mass index ≥35 kg/m² in the year before their procedure; 3) no multiple conflicting bariatric procedure codes on the same day; 4) no revision bariatric procedure code, gastrointestinal cancer diagnosis code, or fundoplasty procedure in the year before the index procedure; 5) no emergency room encounter on the day of index procedure; and 6) a weight measurement 6–18 months after the index procedure. A full description of the design and findings from the weight loss study are available elsewhere.

Exposure

We restricted the analysis to the comparison of Roux-en-Y gastric bypass and sleeve gastrectomy.

Outcome

The outcome of interest was percent total weight loss at 1 year after surgery in comparison with presurgery weight, calculated as [postsurgery weight at 1 year − baseline weight]/baseline weight] and modeled as a continuous variable.

Confounders

We included the same potential confounders identified a priori by the study investigators in the main weight loss analysis based on their subject matter knowledge and input from stakeholders. These confounders included age; sex; race; Hispanic ethnicity; year of procedure; site; combined comorbidity score; baseline weight; the number of days between baseline weight measurement and the index procedure; baseline smoking status; total days of hospitalization in the year prior to surgery; and diagnosis of anxiety, deep vein thrombosis, depression, diabetes, dyslipidemia, eating disorder, gastroesophageal reflux disease, hypertension, infertility, kidney disease, nonalcoholic fatty liver disease, osteoarthritis, polycystic ovarian syndrome, psychosis, pulmonary embolism, sleep apnea, and substance use disorder, 251774 1 / 1 1774 1 / 1
recorded in the year prior to surgery. As described in more
detail below, these confounders were accounted for in the
outcome regression model as individual covariates, propen-
sity scores, or both.

Pooled individual-level data analysis
We used the pooled deidentified individual-level ordinary
least squares regression analysis as the reference analysis.
For this analysis, we wrote distributed programs to run at
each site to pull individual-level data and then combined
all site-specific individual-level data sets together to form a
standard pooled deidentified analysis data set, which included
exposure status, all confounders, site indicators, and the
outcome. To estimate propensity scores within the combined
data set, we fit a logistic regression model, separately for each
site, which predicted the probability of receiving Roux-en-Y
gastric bypass procedure (vs sleeve gastrectomy) using all
the confounders described above as predictors.

We fit three ordinary least squares linear regression
models to examine the difference in percent total weight
loss at 1 year compared with baseline, comparing the Roux-
en-Y gastric bypass procedure with the sleeve gastrectomy
procedure. Model 1 included an exposure indicator variable,
indicator variables for sites, and potential confounders as
individual variables. The purpose of this analysis was to
demonstrate the feasibility and empirical validity of conduct-
ing multivariable-adjusted distributed linear regression in
large real-world distributed data networks such as PCORnet.
Model 2 included the exposure and site indicator variables
and propensity scores (in deciles, defined within each site).
The purpose of this analysis was to examine the feasibility
and empirical validity of combining distributed regression
with propensity scores via modeling. Model 3 included the
exposure and site indicator variables, potential confounders
as individual variables, and propensity scores (in deciles).
This model is an example of a “doubly robust” regression
model – the results would be valid if either the propensity
score model or the outcome model was correctly specified.27,28

Distributed linear regression
Distributed regression is a suite of methods that enable
researchers to conduct multidatabase regression analysis
without the need to centrally combine all individual-level
data from participating sites.16,17 As shown in Figure 1, it
performs the same numeric algorithm as standard regression
that is based on individual-level data but uses only summary
statistics for computation. By following the same computa-
tion process, distributed regression and pooled individual-
level data analysis should theoretically produce statistically
equivalent results.

For distributed linear regression, the total sums of
squares and cross products (SSCP) matrix for the intercept,
dependent variable, and independent variables across all
databases is sufficient to obtain the overall parameter esti-

| ID | E | X1 | X2 | Y       |
|----|---|----|----|--------|
| A001 | 0 | 13.89 | 3.42 | 28.70 |
| A002 | 1 | 18.10 | 1.29 | 27.90 |
| A003 | 0 | 6.41 | 4.86 | 33.10 |
| A004 | 1 | 16.30 | 1.45 | 17.20 |
| A005 | 1 | 17.57 | 2.51 | 21.70 |
| ... |   |     |     | ...    |
| A100 | 0 | 5.78 | 2.53 | 23.76 |

| Type | Name | Intercept | E  | X1  | X2  | Y   |
|------|------|-----------|----|-----|-----|-----|
| SSCP | Intercept | 100.0 | 52.0 | 1,157.1 | 405.9 | 2,235.5 |
| SSCP | E     | 52.0 | 52.0 | 813.2 | 138.1 | 1,060.9 |
| SSCP | X1    | 1,157.1 | 813.2 | 17,751.3 | 3,458.7 | 23,815.8 |
| SSCP | X2    | 405.9 | 138.1 | 3,458.7 | 2,240.8 | 9,573.3 |
| SSCP | Y     | 2,235.5 | 1,060.9 | 23,815.8 | 9,573.3 | 56,911.9 |
| Mean  | 1.0  | 0.5  | 11.6 | 4.1  | 22.4 |
| SD    | 0.0  | 0.5  | 6.6  | 2.5  | 8.4 |
| N     | 100  | 100  | 100  | 100  | 100 |

Variable | Parameter estimate | SE |
----------|--------------------|----|
Intercept | 25.4540 | 3.7969 |
E | -0.4323 | 1.7865 |
X1 | -0.5643 | 0.1432 |
X2 | -0.6564 | 0.4532 |

Figure 1 Computation process of a typical regression analysis.
Note: Numbers are hypothetical and for demonstrative purposes only.
Abbreviations: SSCP, sums of squares and cross products; SE, standard error.
mates and associated covariance matrix. Because of its statistical properties, the total SSCP matrix can be obtained by simply summing up the site-specific SSCP matrices across all databases. In other words, we can request the summary-level SSCP matrix from each participating site and combine these matrices centrally to produce the parameter estimates and SEs (or 95% CIs), as shown in Figure 2. Although standard statistical software packages cannot readily perform distributed linear regression, they can produce the SSCP matrix. For example, SAS software has the procedure PROC REG, which performs linear regression analysis, and it readily inputs and outputs SSCP matrix.

To perform distributed linear regression, we followed a process similar to the one described in the pooled individual-level data analysis, but with more computation done at the site. Figure 3 describes the steps involved in each analysis. Specifically, we modified the distributed program used to pull individual-level data by including additional code that 1) fit a site-specific propensity score model and 2) decomposed the individual-level data into three sets of SSCP matrices, one for each regression model described above. We distributed the modified program to the participating sites for local execution. We then combined these individual site-specific SSCP matrices to produce parameter estimates and SEs.

The distributed SAS packages used to create the site-specific deidentified individual-level data sets (for the reference analysis) and SSCP matrices (for the distributed regression analysis) as well as the SAS programs used by the analysis center to produce the final results are publicly available on GitHub at https://github.com/pcornet-analytics/bariatric.

Comparison

We compared the parameter estimate and standard error (SE) of each variable in each model (Models 1–3) from the pooled individual-level data analysis and distributed regression approach.

![Figure 2 Distributed regression in a multicenter study.](https://www.dovepress.com/)

Note: Numbers are hypothetical and for demonstrative purposes only.

Abbreviation: SSCP, sums of squares and cross products.
Secondary analysis
Due to small sample sizes, we expected some sites to have issues fitting their site-specific propensity score models. Therefore, we reran the analyses excluding sites that had any warnings or errors in their returned SAS logs.

Results
We identified 43,110 eligible patients – 23,963 underwent the Roux-en-Y gastric bypass procedure, and 19,147 patients had the sleeve gastrectomy procedure – from 36 participating data marts. The sample size ranged from 5 to 16,257 across sites. Table 1 summarizes the baseline characteristics of the study cohort by treatment group.

Analyses that included all sites, regardless of warnings or errors when fitting the site-specific propensity score models
For each of the three regression models, the results were identical between the distributed regression analysis and the corresponding pooled individual-level data analysis, as shown in Tables 2–4. Overall, the results showed that Roux-en-Y gastric bypass was associated with greater percent total weight loss than sleeve gastrectomy at 1-year follow-up – approximately 5 percentage points greater in all the three regression models examined. The maximum difference in numerical value was $3.28 \times 10^{-11}$ for the parameter estimate and SE across all the variables in Model 1, $3.24 \times 10^{-11}$ in Model 2, and $3.29 \times 10^{-11}$ in Model 3. Across regression models, adjusting for potential confounders as individual covariates, propensity scores, or both in the models produced similar effect estimates for the comparison of interest.

Analyses that included only sites without any warnings or errors when fitting the site-specific propensity score models
In total, 17 sites successfully executed the SAS package without any warnings or errors in the site-specific propensity score model. There were 36,476 patients – 20,260 Roux-en-Y gastric bypass patients and 16,216 sleeve gastrectomy patients in these 17 sites. The sample size ranged from 242 to 16,257 across these sites. Table S1 summarizes the baseline characteristics of these patients by treatment group. For each of the three regression models, the results were identical between the pooled individual-level data analysis and the distributed regression analysis, as shown in Tables S2–S4. The maximum difference in numerical value was $3.91 \times 10^{-11}$
for the parameter estimate and SE across all the variables in Model 1, $3.35 \times 10^{-11}$ in Model 2, and $3.67 \times 10^{-11}$ in Model 3. Across the regression models, adjusting for potential confounders as individual covariates, propensity scores, or both in the models produced similar effect estimates for the comparison of interest. Not surprisingly, the results from the 17 error-free sites, which comprised 85% (36,476 of 43,110) of the data, were similar to the results obtained from the corresponding analyses that included all data.

**Discussion**

In this study, we demonstrated the feasibility of implementing distributed linear regression in an active, real-world, national distributed data network and the validity of this
analytic method in producing results identical to those obtained from conventional pooled individual-level data analysis for a single continuous outcome. Although some of the prior studies have compared the results from distributed regression with those obtained from the corresponding pooled individual-level data analysis, they were conducted within smaller distributed data networks or more controlled environments (ie, virtual or simulated distributed data networks). The current study is the largest study that demonstrated the feasibility and empirical validity of distributed regression. As these more privacy-protecting methods may be less well understood by researchers, empirical

| Variable          | Parameter estimate | SE       | Parameter estimate | SE       | SE       | SE       |
|-------------------|--------------------|----------|--------------------|----------|----------|----------|
|                   | Pooled individual-level data analysis | Distributed regression | Pooled individual-level data analysis | Distributed regression |
| Exposure          | −0.05312           | −0.05312 | 0.00105            | 0.00105  |
| Age               | −0.01662           | −0.01662 | 0.00106            | 0.00106  |
| 20–44             | 0.01339            | 0.01339  | 0.00218            | 0.00218  |
| 45–64             | 0.02162            | 0.02162  | 0.00133            | 0.00133  |
| Male sex          | 0.00227            | 0.00227  | 0.00147            | 0.00147  |
| White             | Reference          | Reference | Reference          | Reference |
| Black             | 0.02873            | 0.02873  | 0.00130            | 0.00130  |
| Other             | 0.00983            | 0.00983  | 0.00153            | 0.00153  |
| Hispanic ethnicity| 0.00027            | 0.00027  | 0.000147           | 0.000147 |
| Year of procedure |                    |          |                    |          |
| 2005–2009         | −0.00319           | −0.00319 | 0.00218            | 0.00218  |
| 2010              | −0.00301           | −0.00301 | 0.00169            | 0.00169  |
| 2011              | −0.00326           | −0.00326 | 0.00144            | 0.00144  |
| 2012              | Reference          | Reference | Reference          | Reference |
| 2013              | 0.00384            | 0.00384  | 0.00141            | 0.00141  |
| 2014              | 0.00569            | 0.00569  | 0.00145            | 0.00145  |
| 2015              | 0.03664            | 0.03664  | 0.00288            | 0.00288  |
| Comorbidity score | 0.00576            | 0.00576  | 0.00069243         | 0.00069243 |
| Baseline weight   | −0.00025096        | −0.00025096 | 0.00000921        | 0.00000921 |
| Baseline weight proximity | 0.00012330 | 0.00012330 | 0.00001430 | 0.00001430 |
| Smoking           | −0.00657           | −0.00657 | 0.00163            | 0.00163  |
| Days of hospitalization | 0.00017944 | 0.00017944 | 0.00006129 | 0.00006129 |

Notes: Also adjusted for sites (35 indicator variables; results not shown for brevity); Roux-en-Y gastric bypass vs sleeve gastrectomy; measured in the year prior to the surgery; modeled as a continuous variable; number of days between baseline weight measurement and index procedure.

Abbreviations: GERD, gastroesophageal reflux disease; NAFLD, nonalcoholic fatty liver disease; PCORnet, National Patient-Centered Clinical Research Network; PCOS, polycystic ovarian syndrome; SE, standard error.
demonstration of their validity is essential for encouraging their adoption in practice.

To our knowledge, this is also the first study that combined distributed regression with propensity scores to provide even more privacy protection and analytic flexibility than either approach alone. Specifically, the approach first condenses information from a large number of covariates into a single measure and then further processes the condensed individual-level information into highly summarized information.

Propensity scores are a data dimension reduction technique that can summarize a large number of individual covariates into a single, less identifiable scalar. Methods that leverage the property of propensity scores have been shown to produce results identical or highly comparable to those obtained from pooled individual-level data analysis. However, existing propensity score-based approaches only support matched or stratified analysis when sharing summary-level information. In this study, we demonstrated the feasibility of adjusting for propensity scores via regression modeling using only summary-level information in a multicenter study. Existing distributed regression requires all participating sites to fit the same multivariable-adjusted regression model. By combining distributed regression with propensity scores, researchers now have the ability to adjust for different sets of covariates via site-specific propensity score models (more below).

### Transparency and reproducibility in distributed analysis

Without direct access to all the individual-level data from participating sites, researchers are required to specify all the descriptive and inferential analyses a priori as part of the data request process. It is possible and advisable to inspect the data using summary-level information (eg, “Table 1” that describes the patient characteristics at each site) before finalizing the regression model. It is also recommended to include some sites that are able and willing to share individual-level data to test the analytic code before full distribution. As with pooled deidentified individual-level data analysis, it is feasible to fit multiple regression models as sensitivity or secondary analyses, as we did in this study. By being completely explicit about the analysis, the distributed analytic approach minimizes the risk of data dredging and selective reporting. In other words, although the analytic approach used in this study may initially appear opaque to readers who are less familiar with the approach, it is actually more transparent and reproducible.

### Distributed regression for other generalized linear models

Although we only examined distributed linear regression, it is possible to conduct multivariable-adjusted distributed analysis for other commonly used generalized linear models, including logistic, Poisson, and Cox proportional hazards model. Unlike linear regression, which can be completed in a single computation step, the computation process of other regression models requires multiple iterative steps. The iterative process involves exchanging intermediate statistics and interim parameter estimates between the analysis center and participating sites. Manual exchanges of this summary-level information can be too tedious and

---

**Table 3** Results from a linear regression model that adjusted for sites and confounders as propensity score deciles (Model 2) from 36 health systems participating in the PCORnet Bariatric Study

| Variable | Parameter estimate | SE |
|----------|--------------------|----|
| Exposure | Pooled individual-level data analysis | Distributed regression | Pooled individual-level data analysis | Distributed regression |
| PS stratum 1 | Reference | Reference | 0.00113 | 0.00113 |
| PS stratum 2 | -0.00754 | -0.00754 | 0.00209 | 0.00209 |
| PS stratum 3 | -0.00671 | -0.00671 | 0.00210 | 0.00210 |
| PS stratum 4 | -0.00717 | -0.00717 | 0.00211 | 0.00211 |
| PS stratum 5 | 0.00034218 | 0.00034218 | 0.00212 | 0.00212 |
| PS stratum 6 | -0.00583 | -0.00583 | 0.00213 | 0.00213 |
| PS stratum 7 | -0.00135 | -0.00135 | 0.00214 | 0.00214 |
| PS stratum 8 | -0.00435 | -0.00435 | 0.00216 | 0.00216 |
| PS stratum 9 | -0.00523 | -0.00523 | 0.00218 | 0.00218 |
| PS stratum 10 | -0.00812 | -0.00812 | 0.00222 | 0.00222 |

Notes: Also adjusted for sites (35 indicator variables; results not shown for brevity); Roux-en-Y gastric bypass vs sleeve gastrectomy.

Abbreviations: PCORnet, National Patient-Centered Clinical Research Network; PS, propensity score; SE, standard error.
Table 4 Results from a linear regression model that adjusted for sites and confounders as both individual covariates and propensity score deciles (Model 3) from 36 health systems participating in the PCORnet Bariatric study

| Variable                  | Parameter estimate | Distributed regression | SE              | Distributed regression | SE              |
|---------------------------|--------------------|------------------------|-----------------|------------------------|-----------------|
| **Exposure**              |                    |                        |                 |                        |                 |
| Age                       | −0.05355           | −0.05355               | 0.00108         |                        |                 |
| 20–44                     | −0.01668           | −0.01668               | 0.00106         |                        |                 |
| 45–64                     | Reference          | Reference              | Reference       |                        | Reference       |
| 65–80                     | 0.01364            | 0.01364                | 0.00218         |                        | 0.00218         |
| **Male sex**              | 0.02189            | 0.02189                | 0.00134         |                        | 0.00134         |
| **White**                 | Reference          | Reference              | Reference       |                        | Reference       |
| **Race**                  |                    |                        |                 |                        |                 |
| **Hispanic ethnicity**    | 0.00218            | 0.00218                | 0.00148         |                        | 0.00148         |
| **Year of procedure**     |                    |                        |                 |                        |                 |
| 2005–2009                 | −0.00372           | −0.00372               | 0.00241         |                        | 0.00241         |
| 2010                      | −0.00352           | −0.00352               | 0.00187         |                        | 0.00187         |
| 2011                      | −0.00348           | −0.00348               | 0.00146         |                        | 0.00146         |
| 2012                      | Reference          | Reference              | Reference       |                        | Reference       |
| 2013                      | 0.00453            | 0.00453                | 0.00146         |                        | 0.00146         |
| 2014                      | 0.00683            | 0.00683                | 0.00157         |                        | 0.00157         |
| 2015                      | 0.03793            | 0.03793                | 0.00295         |                        | 0.00295         |
| **Comorbidity score**     | 0.00585            | 0.00585                | 0.00069368      |                        | 0.00069368      |
| **Baseline weight**       | −0.00025150        | −0.00025150            | 0.00000922      |                        | 0.00000922      |
| **Baseline weight proximity** | 0.00012628         | 0.00012628             | 0.0001507       |                        | 0.0001507       |
| **Smoking**               | −0.00653           | −0.00653               | 0.00163         |                        | 0.00163         |
| **Days of hospitalization** | 0.00018276         | 0.00018276             | 0.0006130       |                        | 0.0006130       |
| **Diagnosis of**          |                    |                        |                 |                        |                 |
| Anxiety                   | 0.0040061          | 0.0040061              | 0.00119         |                        | 0.00119         |
| Deep vein thrombosis      | 0.00272            | 0.00272                | 0.000530        |                        | 0.000530        |
| Depression                | 0.00403            | 0.00403                | 0.00107         |                        | 0.00107         |
| Diabetes                  | 0.01859            | 0.01859                | 0.00115         |                        | 0.00115         |
| Dyslipidemia              | 0.00164            | 0.00164                | 0.00103         |                        | 0.00103         |
| Eating disorder           | −0.00226           | −0.00226               | 0.00236         |                        | 0.00236         |
| GERD                      | −0.00164           | −0.00164               | 0.00097208      |                        | 0.00097208      |
| Hypertension              | 0.01451            | 0.01451                | 0.00124         |                        | 0.00124         |
| Infertility               | 0.00880            | 0.00880                | 0.00521         |                        | 0.00521         |
| Kidney disease            | 0.00055973         | 0.00055973             | 0.000176        |                        | 0.000176        |
| NAFLD                     | −0.00680           | −0.00680               | 0.00150         |                        | 0.00150         |
| Osseoathritis             | −0.00244           | −0.00244               | 0.00337         |                        | 0.00337         |
| PCOS                      | 0.00127            | 0.00127                | 0.00212         |                        | 0.00212         |
| Psychosis                 | 0.00003263         | 0.00003263             | 0.00226         |                        | 0.00226         |
| Pulmonary embolism        | 0.00757            | 0.00757                | 0.00415         |                        | 0.00415         |
| Sleep apnea               | −0.00177           | −0.00177               | 0.00098965      |                        | 0.00098965      |
| Substance use disorder    | −0.00745           | −0.00745               | 0.00310         |                        | 0.00310         |
| PS stratum 1              | Reference          | Reference              | Reference       |                        | Reference       |
| PS stratum 2              | 0.00180            | 0.00180                | 0.00204         |                        | 0.00204         |
| PS stratum 3              | 0.00362            | 0.00362                | 0.00207         |                        | 0.00207         |
| PS stratum 4              | 0.00058495         | 0.00058495             | 0.00211         |                        | 0.00211         |
| PS stratum 5              | 0.00731            | 0.00731                | 0.00217         |                        | 0.00217         |
| PS stratum 6              | 0.00056111         | 0.00056111             | 0.00223         |                        | 0.00223         |
| PS stratum 7              | 0.00508            | 0.00508                | 0.00231         |                        | 0.00231         |
| PS stratum 8              | 0.00383            | 0.00383                | 0.00242         |                        | 0.00242         |
| PS stratum 9              | 0.00516            | 0.00516                | 0.00257         |                        | 0.00257         |
| PS stratum 10             | 0.00336            | 0.00336                | 0.00285         |                        | 0.00285         |

Notes: *Also adjusted for sites (35 indicator variables; results not shown for brevity); \( ^{1} \) Roux-en-Y gastric bypass vs sleeve gastrectomy; \( ^{2} \) measured in the year prior to the surgery; \( ^{3} \) modeled as a continuous variable; \( ^{4} \) number of days between baseline weight measurement and index procedure.

Abbreviations: GERD, gastroesophageal reflux disease; NAFLD, nonalcoholic fatty liver disease; PCORnet, National Patient-Centered Clinical Research Network; PCOS, polycystic ovarian syndrome; PS, propensity score; SE, standard error.
labor-intensive to be practical in actual multicenter studies. However, there are a number of statistical packages and stand-alone software that enable researchers to perform distributed regression and partially or fully automate the file transfer process.\textsuperscript{19,20,31–33} We chose not to assess the operational performance (eg, runtime) of our distributed regression analysis because it is highly network-dependent. Some multicenter studies have better (or less ideal) technical infrastructure to allow distributed regression to be conducted more (or less) efficiently.

Site-specific propensity score models vs a “global” propensity score model

We fit site-specific propensity score models because propensity scores are sensitive to the prevalence of exposure in the population under study. As the prevalence of bariatric procedures might vary across databases due to differences in clinical practice and other factors, patients from different sites might not be comparable even if they had the same propensity score value. In addition, a given covariate (eg, diabetes) may have different effects on the probability of receiving a specific bariatric procedure at different sites. Therefore, it is generally advisable to estimate propensity scores by site and perform the analysis accounting for data source, eg, match on propensity scores within site, stratify jointly on site-specific propensity scores and site, and include both propensity scores and site as regressors.

However, it is not always possible to estimate propensity scores by site because some sites might be too small to accommodate a robust propensity score model. Indeed, the number of patients in some sites in the current study was too low to allow the conventional rule of requiring 7–10 “outcome events” (ie, exposed patients) per covariate in the propensity score model.\textsuperscript{14,35} In a distributed environment, researchers have the ability to inspect the log or other summary-level output (eg, a patient characteristic table) to assess the feasibility of fitting site-specific propensity score models.

An alternative is to fit one propensity score model that adjusted for the confounders, site, and additional interaction terms between confounders and site, which was done in the primary analysis for this study.\textsuperscript{25} This allows sites that do not have sufficient data to fit site-specific propensity score models to contribute their data to the analysis, with an additional assumption that the global propensity score model is correctly specified. It is straightforward to fit such a model in the pooled individual-level data analysis. Although we did not do it here, in principle it would be possible to fit a global propensity score model using distributed logistic regression. This is a topic for future research.

Some sites may have additional confounder information that can be included in the study. For example, some participating health systems in this study may have dietary information from patients who underwent bariatric procedure. For simplicity, we only adjusted for confounders that were available in all participating sites. It is straightforward computationally to include different sets of confounders in the site-specific propensity score models, but the use of different confounders across sites would need to be justified scientifically. Future studies should examine the strengths and limitations of allowing site-specific confounder lists.

Limitations of distributed regression approach

Although we were able to illustrate the statistical equivalence of the distributed linear regression approach to pooled individual-level data analysis for a one-time continuous outcome, we were not able to fully replicate the primary analysis in the main study, which used a linear mixed-effects model to simultaneously estimate the association between bariatric surgeries and percent total weight loss at 1, 3, and 5 years postsurgery.\textsuperscript{25} In the main study, the estimated difference in percent weight loss at 1 year postsurgery was −5.9% comparing Roux-en-Y gastric bypass with sleeve gastrectomy. The results from the distributed linear regression analyses, which were between −5.3% and −5.5% across the three models, were comparable. We may have observed greater dissimilarities if we estimated a 3- or 5-year postsurgery difference where the advantages of using repeated measure analysis with a linear mixed-effects model are more pronounced due to missing data assumptions being relaxed (eg, mixed-effects models assume that data are missing at random borrowing from more outcomes being measured). Distributed regression has not been fully developed for models that examine repeated outcomes or repeated exposures.

In addition, certain model diagnostics cannot currently be performed with summary-level information. For example, creating residual plots to examine the normality assumption in linear regression requires sharing of individual-level data. It is possible to introduce noise to the individual-level data but doing so could impact its fidelity.\textsuperscript{36} It is also more challenging to identify or assess idiosyncratic data issues at participating sites without direct access to all individual-level data being analyzed. Enhancing the capability of distributed regression to perform other tasks that can readily
be performed with individual-level data is another topic for future research.

**Other privacy-protecting analytic and data-sharing methods**

Prior studies have demonstrated the feasibility and validity of using summary-level risk-set data to perform propensity score-matched and propensity score-stratified analysis.\(^{11-13}\) Recent work has explored the feasibility of conducting inverse probability-weighted analysis using risk-set data.\(^{37}\) Our study shows the feasibility of adjusting for propensity scores via modeling using only summary-level information in multicenter studies. In addition to summary score-based methods and distributed regression, meta-analysis of site-specific estimates is another analytic option that does not require sharing of individual-level data in a multicenter study. Studies have shown that meta-analysis produced results comparable to those from pooled individual-level data analysis.\(^{11-15,38}\) However, unlike distributed regression, meta-analysis generally produces similar, but not identical, results to those from pooled individual-level data analysis.

Instead of using distributed regression, it would also be possible to request stratified counts of all unique combinations of the confounders from participating sites to fit a regression model in a multicenter study. The approach requires all confounders to be binary or categorical variables.\(^{39}\) It would be feasible to employ that approach for Model 2, but not the other two models examined in the study because of the large number of confounders in these two models and because some of them were modeled as continuous variables.

**Conclusion**

We demonstrated the feasibility and empirical validity of performing distributed linear regression analysis for one-time linear outcomes within a large distributed data network and the feasibility and validity of combining distributed regression and propensity scores to offer additional privacy protection and analytic flexibility. Along with other existing privacy-protecting analytic methods, distributed regression is a viable and valid analytic option that allows researchers to analyze data that may otherwise be inaccessible.

**Availability of data and computing code**

The analytic code is available at https://github.com/pcornet-analytics/bariatric.

**Acknowledgments**

The PCORnet Bariatric Study Collaborative includes key investigators and stakeholders who made substantial contributions to the manuscript as authors. The PCORnet Bariatric Study Collaborative authorship list includes: Sengwee Toh (Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute), Robert Wellman (Kaiser Permanente Washington Health Research Institute), R Yates Coley (Kaiser Permanente Washington Health Research Institute), Casie Horgan (Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute), Jessica Sturtevant (Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute), Eric Moyneur (Statlog Econometrics, Inc.), Cheri Janning (Duke Clinical and Translational Science Institute), Roy Pardee (Kaiser Permanente Washington Health Research Institute), Karen J Coleman (Kaiser Permanente Southern California), David Arterburn (Kaiser Permanente Washington Health Research Institute), Kathleen McTigue (Department of Medicine, University of Pittsburgh), Jane Anau (Kaiser Permanente Washington Health Research Institute), Andrea J Cook (Kaiser Permanente Washington Health Research Institute); Joseph Vitello (Jesse Brown VA Medical Center); Elisha Malanga (COPD Foundation); Corrigan L McBride (Department of Surgery, University of Nebraska Medical Center); James McClay (University of Nebraska Medical Center); David Schlundt (Vanderbilt University); Cynthia Blalock (Vanderbilt University Medical Center); Tammy St Clair (patient partner); Joseph Nadglowski, Jr (Obesity Action Coalition); Anita P Courcoulas (University of Pittsburgh Medical Center); Marc Michalsky (Nationwide Children’s Hospital); Nirav Desai (Division of Gastroenterology, Hepatology, and Nutrition, Boston Children’s Hospital); Alberto Odor (University of California, Davis); William Richardson (Ochsner Clinic); Caroline M Apovian and William G Adams (Boston University School of Medicine/Boston Medical Center); Elizabeth Cirelli and Ali Tavakkoli (Brigham and Women’s Hospital); Roni Zeiger (Smart Patients, Inc.); Laura Rasmussen-Torvik (Northwestern University Feinberg School of Medicine); Gabrielle Purcell and Jeffrey Brown (Department of Population Medicine, Harvard Medical School, Harvard Pilgrim Health Care Institute); John H Holmes (Perelman School of Medicine, University of Pennsylvania); Ana Emiliano (The Rockefeller University); Jonathan N’Tobin (Clinical Directors Network and The Rockefeller University); Jeanne M Clark (Johns Hopkins University and Health Plan); Wendy Bennett (The Johns Hopkins University...
the following individuals for their invaluable contributions to PCORnet and the PCORnet Bariatric Study: Stephen R Perry, Kin Lam, David Hawkes, Thomas Dunod, and Kelli Kinsman (Kaiser Permanente Washington Health Research Institute); Shelly Sital (The Chicago Community Trust); Elizabeth Tarlov (University of Illinois at Chicago); Jasmin Phua (Medical Research Analytics and Informatics Alliance); Mia Gallagher, Lindsey Petro, and Beth Syat (Harvard Pilgrim Health Care Institute and Harvard Medical School); Prakash Nadkarni and Elizabeth Chrischilles (University of Iowa); Steffani Roush, Robert Greenlee, and Laurel Verhagen (Marshfield Clinic Research Institute); Umberto Tassignard (University of Wisconsin); Phillip Reeder, Shibly Antony, Rania AlShahrouri (University of Texas – Southwestern Medical Center); James Campbell, Russell Buzalko, and Jay Pedersen (University of Nebraska Medical Center); Russel Waitman and Dan Connolly (University of Kansas Medical Center); Russel Rothman, David Crenshaw, and Katie Worley (Vanderbilt University Medical Center); Emily Pfaff, Robert Bradford, Kelly Walters, Tim Carey, Timothy Farrell, and D Wayne Overby (University of North Carolina); Majja Neville-Williams and Rhonda G Kost (The Rockefeller University); Elizabeth Shenkman, William Hogan, Kathryn McAuliffe, and Gigi Lipori (University of Florida); Rebecca Zuvich Essner (Florida Hospital); Howard Su, Michael George, Michael J Becich, Barbara Postol, Giselle G Hamad, Ramesh C Ramanathan, Bestoun H Ahmed, William F Gourash, Bill Shirley, Chuck Borromeo, and Desheng Li (University of Pittsburg); Anthony T Petrick, Ilene Ladd, Preston Biegley, H Lester Kirchner (Geisinger); Daniel E Ford, Michael A Schweitzer, Karl Burke, Harold Lehmann, Megan E Gauvey-Kern, and Diana Gumus (Johns Hopkins); Rachel Hess, Meghan Lynch, and Reid Holbrook (University of Utah); Jody McCullough, Matt Bolton, Wenke Hwang, Cynthia Chuang, Ann Rogers, and Alison Bower (Penn State); Cecilia Dobi, Michael Edwards, Mark Weiner, Anuradha Paranjape, Sharon J Herring, and Patricia Bernard (Temple University); Janet Zahner, Parth Divekar, Keith Marsolo, and Lisa Boerger (Cincinnati Children’s Hospital); Kimberly J Holmquist (Kaiser Permanente Southern California); Ray Pablo, Roni Bracha, and Robynn Zender (University of California at Irvine); Lucila Olno-Machado, Paulina Paul, and Michele Day (University of California, San Diego); Thomas Carton, Elizabeth Crull, and Iben McCormick-Rickett (Louisiana Public Health Institute); Ashley Vernon, Malcom Robinson, Scott Shikora, David Spector, Eric Sheu, Edward Mun, Matthew Hutter, Shawn Murphy, Jeffrey Klann, and Denise Gee (Partners Healthcare); Daniel Jones, Benjamin Schneider, Griffin Weber, and Robert Andrews (Beth Israel Deaconess Medical Center); Thomas H Inge (Children’s Hospital Colorado and University of Colorado, Denver); Lydia Bazzano (Tulane University School of Public Health and Tropical Medicine); Elizabeth Nauman (Louisiana Public Health Institute); Denise M Hynes (US Department of Veterans Affairs and Oregon State University); David Meltzer (The University of Chicago); Bipan Chand (Loyola University Medical Center); Jeffrey J VanWormer (Marshfield Clinic Research Institute); Lindsay G Cowell (UT Southwestern Medical Center); Lawrence P Hanrahan (University of Wisconsin – Madison); Meredith Duke (University of North Carolina); Daniel M Herron (Icahn School of Medicine at Mount Sinai); Sameer Malhotra (Weill Cornell Medicine); Jiang Bian (University of Florida); Michelle R Lent (Geisinger); Jennifer L Kraschnewski and Julie Tice (Penn State College of Medicine); Michael A Edwards (Lewis Katz School of Medicine at Temple University); Molly M Conroy (Division of General Internal Medicine, University of Utah); Matthew F Daley (Institute for Health Research, Kaiser Permanente Colorado); Michael Horberg (Kaiser Permanente Mid-Atlantic Permanente Research Institute); Jay R Desai (HealthPartners Institute); Stephanie L Fitzpatrick (Kaiser Permanente Center for Health Research); Douglas S Bell (Department of Medicine, David Geffen School of Medicine at UCLA, and UCLA Clinical and Translational Science Institute); Erin Roe (Baylor Scott & White Endocenter Center); Ellen P McCarthy (Beth Israel Deaconess Medical Center); Andrew O Odegard (Department of Epidemiology, School of Medicine, University of California, Irvine); Neely A Williams (Community Partners’ Network, Inc.); and Sameer Murali (Southern California Permanente Medical Group).

The study was funded by PCORI through PCORI Award OBS-1505-30683. The views expressed in this manuscript are solely those of the authors and do not reflect the views of PCORnet® or PCORI. The authors would like to acknowledge the following individuals for their invaluable contributions to PCORnet and the PCORnet Bariatric Study: Stephen R Perry, Kin Lam, David Hawkes, Thomas Dunod, and Kelli Kinsman (Kaiser Permanente Washington Health Research Institute); Shelly Sital (The Chicago Community Trust); Elizabeth Tarlov (University of Illinois at Chicago); Jasmin Phua (Medical Research Analytics and Informatics Alliance); Mia Gallagher, Lindsey Petro, and Beth Syat (Harvard Pilgrim Health Care Institute and Harvard Medical School); Prakash Nadkarni and Elizabeth Chrischilles (University of Iowa); Steffani Roush, Robert Greenlee, and Laurel Verhagen (Marshfield Clinic Research Institute); Umberto Tassignard (University of Wisconsin); Phillip Reeder, Shibly Antony, Rania AlShahrouri (University of Texas – Southwestern Medical Center); James Campbell, Russell Buzalko, and Jay Pedersen (University of Nebraska Medical Center); Russel Waitman and Dan Connolly (University of Kansas Medical Center); Russel Rothman, David Crenshaw, and Katie Worley (Vanderbilt University Medical Center); Emily Pfaff, Robert Bradford, Kelly Walters, Tim Carey, Timothy Farrell, and D Wayne Overby (University of North Carolina); Majja Neville-Williams and Rhonda G Kost (The Rockefeller University); Elizabeth Shenkman, William Hogan, Kathryn McAuliffe, and Gigi Lipori (University of Florida); Rebecca Zuvich Essner (Florida Hospital); Howard Su, Michael George, Michael J Becich, Barbara Postol, Giselle G Hamad, Ramesh C Ramanathan, Bestoun H Ahmed, William F Gourash, Bill Shirley, Chuck Borromeo, and Desheng Li (University of Pittsburg); Anthony T Petrick, Ilene Ladd, Preston Biegley, H Lester Kirchner (Geisinger); Daniel E Ford, Michael A Schweitzer, Karl Burke, Harold Lehmann, Megan E Gauvey-Kern, and Diana Gumus (Johns Hopkins); Rachel Hess, Meghan Lynch, and Reid Holbrook (University of Utah); Jody McCullough, Matt Bolton, Wenke Hwang, Cynthia Chuang, Ann Rogers, and Alison Bower (Penn State); Cecilia Dobi, Michael Edwards, Mark Weiner, Anuradha Paranjape, Sharon J Herring, and Patricia Bernard (Temple University); Janet Zahner, Parth Divekar, Keith Marsolo, and Lisa Boerger (Cincinnati Children’s Hospital); Kimberly J Holmquist (Kaiser Permanente Southern California); Ray Pablo, Roni Bracha, and Robynn Zender (University of California at Irvine); Lucila Olno-Machado, Paulina Paul, and Michele Day (University of California, San Diego); Thomas Carton, Elizabeth Crull, and Iben McCormick-Rickett (Louisiana Public Health Institute); Ashley Vernon, Malcom Robinson, Scott Shikora, David Spector, Eric Sheu, Edward Mun, Matthew Hutter, Shawn Murphy, Jeffrey Klann, and Denise Gee (Partners Healthcare); Daniel Jones, Benjamin Schneider, Griffin Weber, and Robert Andrews (Beth Israel Deaconess Medical Center); Thomas H Inge (Children’s Hospital Colorado and University of Colorado, Denver); Lydia Bazzano (Tulane University School of Public Health and Tropical Medicine); Elizabeth Nauman (Louisiana Public Health Institute); Denise M Hynes (US Department of Veterans Affairs and Oregon State University); David Meltzer (The University of Chicago); Bipan Chand (Loyola University Medical Center); Jeffrey J VanWormer (Marshfield Clinic Research Institute); Lindsay G Cowell (UT Southwestern Medical Center); Lawrence P Hanrahan (University of Wisconsin – Madison); Meredith Duke (University of North Carolina); Daniel M Herron (Icahn School of Medicine at Mount Sinai); Sameer Malhotra (Weill Cornell Medicine); Jiang Bian (University of Florida); Michelle R Lent (Geisinger); Jennifer L Kraschnewski and Julie Tice (Penn State College of Medicine); Michael A Edwards (Lewis Katz School of Medicine at Temple University); Molly M Conroy (Division of General Internal Medicine, University of Utah); Matthew F Daley (Institute for Health Research, Kaiser Permanente Colorado); Michael Horberg (Kaiser Permanente Mid-Atlantic Permanente Research Institute); Jay R Desai (HealthPartners Institute); Stephanie L Fitzpatrick (Kaiser Permanente Center for Health Research); Douglas S Bell (Department of Medicine, David Geffen School of Medicine at UCLA, and UCLA Clinical and Translational Science Institute); Erin Roe (Baylor Scott & White Endocenter Center); Ellen P McCarthy (Beth Israel Deaconess Medical Center); Andrew O Odegard (Department of Epidemiology, School of Medicine, University of California, Irvine); Neely A Williams (Community Partners’ Network, Inc.); and Sameer Murali (Southern California Permanente Medical Group).
Medical Center); Donald Hess, Brian Carmine, Miguel Burch, and Galina Lozinski (Boston Medical Center); Ken Mandl, Jessica Lyons, and Margaret Vella (Harvard Medical School); and Joseph Skelton and Kun Wei (Wake Forest Integrated Health System).

Disclosure
Dr Arterburn reports NIH funding outside the submitted work. Mr Moyneur reports StatLog was paid consulting fees to conduct study programming. All authors received funding from PCORI to support the submitted work. The authors report no other conflicts of interest in this work.

References
1. Behrmann RE, Benner JS, Brown JS, McClellan M, Woodcock J, Platt R. Developing the Sentinel System—a national resource for evidence development. N Engl J Med. 2011;364(6):498–499.
2. Collins FS, Hudson KL, Briggs JP, Lauer MS. PCORnet: turning a dream into reality. J Am Med Inform Assoc. 2014;21(4):576–577.
3. Richesson RL, Hammond WE, Nahm M, et al. Electronic health records based phenotyping in next-generation clinical trials: a perspective from the NIH Health Care Systems Collaboratory. J Am Med Inform Assoc. 2013;20(2):c226–c231.
4. Sankar PL, Parker LS. The Precision Medicine Initiative’s All of Us Research Program: an agenda for research on its ethical, legal, and social issues. Genet Med. 2017;19(7):743–750.
5. Sussa S, Henry D, Caetano R, et al; Canadian Network for Observational Drug Effect Studies (CNODES). CNODES: the Canadian network for Observational Drug Effect Studies. Open Med. 2012;6(4):e134–e140.
6. Reynolds RF, Kurz X, de Groot MC, et al. Electronic health records based phenotyping in next-generation clinical trials: a perspective from the NIH Health Care Systems Collaboratory. J Am Med Inform Assoc. 2013;20(2):c226–c231.
7. Simon GE, Coronado G, DeBar LL, et al. Data Sharing and Embedded Research. Ann Intern Med. 2017;167(9):668–670.
8. Brown JS, Holmes JH, Shah K, Hall K, Lazarus R, Platt R. Distributed health data networks: a practical and preferred approach to multi-institutional evaluations of comparative effectiveness, safety, and quality of care. Med Care. 2010;48(6 Suppl 1):S45–S51.
9. Toh S, Platt R, Steiner JF, Brown JS. Comparative-effectiveness research in distributed health data networks. Clin Pharmacol Ther. 2011;90(6):883–887.
10. Mazor KM, Richards A, Gallagher M, et al. Stakeholders’ views on data sharing in multicenter studies. J Comp Eff Res. 2017;6(6):537–547.
11. Rassen JA, Avorn J, Schneeweiss S. Multivariate-adjusted pharmacoepidemiologic analyses of confidential information pooled from multiple health care utilization databases. Pharmacoepidemiol Drug Saf. 2010;19(8):848–857.
12. Toh S, Gagne JJ, Rassen JA, Fireman BH, Kulldorff M, Brown JS. Confounding adjustment in comparative effectiveness research conducted within distributed research networks. Med Care. 2013;51(8 Suppl 3):S4–S10.
13. Toh S, Shetty S, Powers JD, Arterburn D. Privacy-preserving analytical methods for multisite comparative effectiveness and patient-centered outcomes research. Med Care. 2014;52(7):664–668.
14. Toh S, Reichman ME, Houstoun M, et al. Multivariable confounding adjustment in distributed data networks without sharing of patient-level data. Pharmacoepidemiol Drug Saf. 2013;22(11):1171–1177.
15. Platt RW, Dormuth CR, Chateau D, Filion K. Observational Studies of Drug Safety in Multi-Database Studies: Methodological Challenges and Opportunities. EGEMS (Wash DC). 2016;4(1):1221.
16. Karr AF, Lin X, Sanil AP, Reiter JP. Secure Regression on Distributed Databases. J Comput Graph Stat. 2005;14(2):1–18.
17. Fienberg SE, Fulp WJ, Slavković AB, Wrobel TA. “Secure” log-linear and logistic regression analysis of distributed databases. Lect Notes Comput Sci. 2006;4032:277–290.
18. Wolfson M, Wallace SE, Masca N, et al. DataSHIELD: resolving a conflict in contemporary bioscience—performing a pooled analysis of individual-level data without sharing the data. Int J Epidemiol. 2010;39(5):1372–1382.
19. El Emam K, Samet S, Arbuckle L, Tamblyn R, Earle C, Kantarcioğlu M. A secure distributed logistic regression protocol for the detection of rare adverse drug events. J Am Med Inform Assoc. 2013;20(3):453–461.
20. Wu Y, Jiang X, Kim J, Ohno-Machado L. Grid Binary Logistic Regression (GLORE): building shared models without sharing data. J Am Med Inform Assoc. 2012;19(5):758–764.
21. Meeker D, Jiang X, Matheny ME, et al. A system to build distributed multivariate models and manage disparate data sharing policies: implementation in the scalable national network for effectiveness research. J Am Med Inform Assoc. 2015;22(6):1187–1195.
22. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. J Am Med Inform Assoc. 2014;21(4):578–582.
23. PCORnet Common Data Model (CDM). Available from: http://www.pcornet.org/pcornet-common-data-model/. Accessed April 26, 2018.
24. Toh S, Rasmussen-Torvik LJ, Harmata EE, et al. The National Patient-Centered Clinical Research Network (PCORnet) Bariatric Study Cohort: Rationale, Methods, and Baseline Characteristics. JMIR Res Protoc. 2017;6(12):e222.
25. Arterburn D, Wellman R, Emiliano A, et al. Comparative effectiveness and safety of bariatric procedures for weight loss: A PCORnet cohort study. Ann Intern Med. Epub 2018 Oct 31.
26. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol. 2011;64(7):749–759.
27. Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. Stat Med. 2004;23(19):2937–2960.
28. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. Biometrika. 2005;91(4):903–923.
29. Fienberg SE, Karr AF, Nardi Y, Slavković AB. Secure logistic regression with multi-party distributed databases. Proceedings of the 56th Session of the ISI. The Bulletin of the International Statistical Institute; Alexandria, VA, 2007: 3506–3513.
30. Slavković AB, Nardi Y, Tibbits MM. Secure logistic regression of horizontally and vertically partitioned distributed databases. Proceedings of Workshop on Privacy and Security Aspects of Data Mining. Washington, DC: IEEE Computer Society Press; 2007: 723–728.
31. Lu CL, Wang S, Ji Z, et al. WebDiSCO: a web service for distributed cox model learning without patient-level data sharing. J Am Med Inform Assoc. 2015;22(6):1212–1219.
32. Gaye A, Marcon Y, Isaeva J, et al. DataSHIELD: taking the analysis to the data, not the data to the analysis. Int J Epidemiol. 2014;43(6):1929–1944.
33. Her QL, Malenfant JM, Malek S, et al. A query workflow design to co-phenotyping and statistical matching in co-epidemiologic analyses of confidential information pooled from multiple health care utilization databases. J Clin Epidemiol. 2003;158(3):280–287.
34. Cepeda MS, Boston R, Farrar JT, Strom BL. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. Am J Epidemiol. 2003;158(3):280–287.
35. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the performance of logistic regression quantiles for developing a classification rule when data sets are limited. J Clin Epidemiol. 1996;49(12):1373–1379.
36. Kargupta H, Datta S, Wang Q, Sivakumar K. On the privacy preserving properties of random data perturbation techniques. Paper presented at: Third IEEE International Conference on Data Mining 2003; Melbourne, Florida.
37. Yoshida K, Gruber S, Fireman BH, Toh S. Comparison of privacy-protecting analytic and data-sharing methods: A simulation study. *Pharmacoepidemiol Drug Saf*. 2018;27(9):1034–1041.

38. Rassen JA, Solomon DH, Curtis JR, Herrinton L, Schneeweiss S. Privacy-maintaining propensity score-based pooling of multiple databases applied to a study of biologics. *Med Care*. 2010;48(6 Suppl):S83–S89.

39. Cook AJ, Wellman RD, Shoaibi A, Nelson JC. Feasibility of new Mini-Sentinel group sequential monitoring methods in a distributed setting – Implementation in practice. Available from: https://www.sentinelinitiative.org/sites/default/files/Methods/Mini-Sentinel_New-Mini-Sentinel-Group-Sequential-Methods_Distributed-Data-Setting_0.pdf. Accessed January 12, 2018.