`

OHDSI METHODS LIBRARY:

SOFTWARE VALIDITY

&

REGULATORY REQUIREMENTS

\*\*\* draft for community review \*\*\*

March 20, 2018

Observational Health Data Science and Informatics

New York, New York

Email: schuemie@ohdsi.org

Contents

1. [**Introduction**](#_gjdgxs) **3**

2. [**Observational Health Data Science and Informatics (OHDSI)**](#_1fob9te) **4**

[Mission, Vision & Values](#_mbbr7dxd142g) 4

[Organization](#_vxaw5kkuy44f) 4

3. [**The OHDSI Methods Library**](#_3znysh7) **5**

4. [**Regulations and guidelines pertaining to the Methods Library**](#_vvt1tqdjoeru) **7**

[Regulations and guidelines for software used in RCTs](#_z3n55457lf4r) 7

[Regulations and guidelines for software used in observational studies](#_20aliig0034c) 7

[Differences between RCTs and observational studies](#_5i8ofebqeide) 7

[Requirements for the OHSDI Methods Library](#_vhxjgzvnhkyb) 8

5. [**Software Development Process**](#_1t3h5sf) **9**

[Source Code Management](#_2s8eyo1) 9

[Documentation](#_17dp8vu) 10

[Availability of Current and Historical Archive Versions](#_26in1rg) 10

[Maintenance, Support and Retirement](#_lnxbz9) 10

[Qualified Personnel](#_35nkun2) 10

[Physical and Logical Security](#_1ksv4uv) 10

[Disaster Recovery](#_44sinio) 10

6. [**Testing**](#_srbuyq8sx2u6) **11**

[Unit test](#_vraeh5lcgd8l) 11

[Method tests for population-level effect estimation](#_6kysqpjwjhbg) 11

7. [**Conclusions**](#_udsf0khd5u18) **12**

8. [**References**](#_osv9seyvu2t4) **13**

[**Appendix A: Method Library packages**](#_ya3y86t6w8y8) **14**

[CaseControl](#_3dcbw97drm6r) 14

[CaseCrossover](#_7ndthcfkkuos) 14

[CohortMethod](#_xwua1re4c4do) 14

[Cyclops](#_ygyo9jkaozdg) 15

[DatabaseConnector](#_pqu5yyx4jr3w) 15

[EmpiricalCalibration](#_wfvxxui4jgyh) 16

[EvidenceSynthesis](#_td31cvdrcgak) 16

[FeatureExtraction](#_qoluj78rxenx) 17

[MethodEvaluation](#_i7s6t9abstm4) 17

[OhdsiRTools](#_qrgzqwq7aoip) 18

[PatientLevelPrediction](#_z5jmnygnhtj9) 18

[SelfControlledCaseSeries](#_zcljf7nh58d6) 18

[SelfControlledCohort](#_8jmzhh75p9ss) 19

[SqlRender](#_kbn6rkryn0ya) 19

[**Appendix B: Results of the method evaluation**](#_jzllf5ewtdl2) **20**

# 

# Introduction

Randomized clinical trials (RCTs) serve as the cornerstone for causal evidence about medical products, but evidence from these trials may be limited by an insufficient number of persons exposed, insufficient length of exposure, and inadequate coverage of the target population, factors that limit external generalizability and ethical impediments when trying to randomize patients into arms with potentially harmful interventions. Observational studies can complement RCTs by testing clinical hypotheses on existing clinical, or observational, data using methods that correct for non-random treatment assignment as part of the effect estimation or hypothesis testing process. Today, such observational studies are conducted by pharmaceutical manufacturers, academic and commercial research organizations and individual clinical investigators.

The validity of the evidence from observational research depends on many factors, including the quality of the data, presence of measured or unmeasured confounding, the study design used, and accurate and adequate reporting. Here we focus on a single aspect: the validity of the analysis software. We aim to answer the question “does the analysis code do what it is supposed to do?”. Specifically, we focus on a set of software that is reusable across observational studies: the Observational Health Data Science and Informatics (OHDSI) Methods Library.

We first describe OHDSI (section 2) and the OHDSI Methods Library (section 3), thereby establishing the scope of any statements made in this document. We then review regulatory requirements for software used in RCTs and software used in observational studies and our interpretation of these requirements as they pertain to the OHDSI Methods Library (section 4). We then describe how the OHDSI Methods Library currently meets those requirements (sections 5 and 6). We end with some concluding thoughts (section 7).

This document is not intended to be prescriptive, does not render a legal opinion and does not confer or impart any binding or other legal obligation. It should be utilized by the reader and his or her organization as one component in the process of making informed decisions as to how best to meet relevant obligations within their own professional working environment.

**OHDSI makes no warranties, expressed or implied, in this document.**

# Observational Health Data Science and Informatics (OHDSI)

OHDSI is a multi-stakeholder, interdisciplinary collaborative that strives to bring out the value of observational health data through large-scale analytics. Its research community enables active engagement across multiple disciplines (e.g., clinical medicine, biostatistics, computer science, epidemiology, life sciences) and spans multiple stakeholder groups (e.g., researchers, patients, providers, payers, product manufacturers, regulators). OHDSI has established an international network of researchers and observational health databases with a central coordinating center housed at Columbia University.

## Mission, Vision & Values

OHDSI’s Mission:

* To improve health, by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care.

OHDSI’s Vision:

* A world in which observational research produces a comprehensive understanding of health and disease.

OHDSI’s Objectives:

* Innovation: Observational research is a field which will benefit greatly from disruptive thinking. OHDSI actively seek and encourage fresh methodological approaches in its work.
* Reproducibility: Accurate, reproducible, and well-calibrated evidence is necessary for health improvement.
* Community: Everyone is welcome to actively participate in OHDSI, whether they are a patient, a health professional, a researcher, or someone who simply believes in OHDSI’s cause.
* Collaboration: OHDSI works collectively to prioritize and address the real world needs of the OHDSI community’s participants.
* Openness: OHDSI strives to make all its community’s proceeds open and publicly accessible, including the methods, tools and the evidence that OHDSI generates.
* Beneficence: OHDSI seeks to protect the rights of individuals and organizations within its community at all times.

## Organization

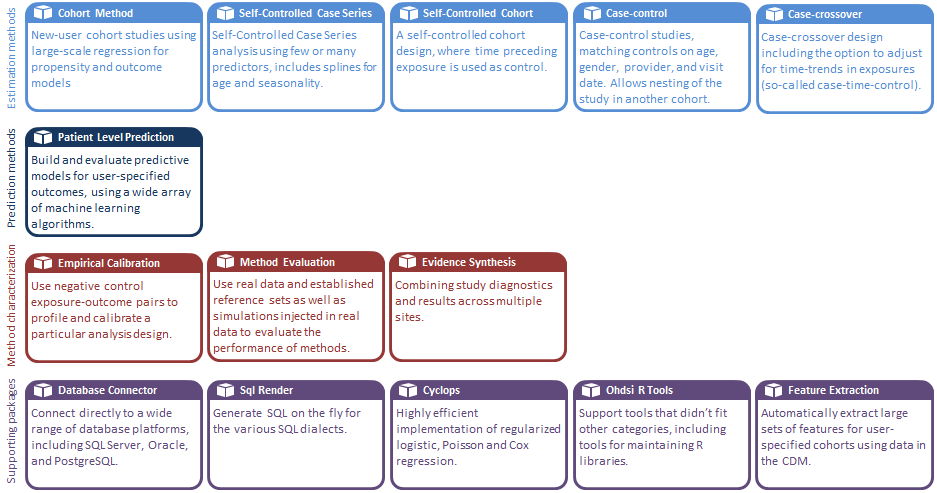
OHDSI has a central coordinating center, but is otherwise decentralized. OHDSI has several workgroups that have taken specific tasks onto themselves. Workgroups are open to all, but have non-rotating leaderships. Especially relevant for this document are the OHDSI Population-Level Estimation Workgroup and the OHDSI Patient-Level Prediction Workgroup, which together are responsible for developing and maintaining the OHDSI Methods Library. The OHDSI Population-Level Estimation Workgroup is headed by Drs. Marc Suchard and Martijn Schuemie. The OHDSI Patient-Level Prediction Workgroup his headed by Drs. Peter Rijnbeek and Jenna Reps.

# 

# The OHDSI Methods Library

The OHDSI Methods Library is a set of open source R packages for population-level effect estimation and patient-level prediction. Several observational study designs are implemented for population-level effect estimation: the new-user cohort design (optionally using propensity scores), self-controlled case series, case-control, case-crossover, and self-controlled cohort. Patient-level prediction follows a single design framework which is flexible in terms of its constituent components such as the machine-learning algorithm used (e.g. deep learning or regularized regression). The packages offer R functions that together can be used to perform an observation study from data to models, estimates and supporting statistics, figures, and tables. The packages interact directly with observational data in the Common Data Model (CDM) 9, and are designed to support both large datasets and large numbers of analyses (e.g. for testing many hypotheses including control hypotheses, and testing many analyses design variations). For this purpose, each Method package includes functions for specifying and subsequently executing multiple analyses efficiently. The Methods Library support best practices for use of observational data as learned from previous and ongoing research, such as transparency, reproducibility, as well as measuring of the operating characteristics of methods in a particular context and subsequent empirical calibration of estimates produced by the methods.

At this moment, “The OHDSI Methods Library” includes the following packages:



**Figure 1.** Packages included in the OHDSI Methods Library.

As shown in Figure 1, the Methods Library not only includes packages specific to the estimation and prediction methods, but also several supporting packages such as packages supporting communication with the database containing the data in CDM format (DatabaseConnector and SqlRender) as well as a package for large scale regularized regression (Cyclops).

This document is not in any fashion applicable to other OHDSI software made available via other parties, such as users or even members of OHDSI who may, from time to time, make their software available via the OHDSI GitHub repositories. The set of packages included in the OHDSI Methods Library, as depicted in Figure 1, may change over time.

# Regulations and guidelines pertaining to the Methods Library

Here we review the most relevant regulations and guidelines for software used in RCTs and observational studies.

## Regulations and guidelines for software used in RCTs

The use of software for the collection, manipulation, analysis and presentation of RCT data is regulated in great detail by two principal regulatory entities: The United States Food and Drug Administration (FDA) and the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals in Human Use (ICH). Similar governmental and regulatory bodies in the international community, such as the European Medicines Agency (EMA) and the Japanese Pharmaceuticals and Medical Devices Agency (PMDA) oversee activities within their respective jurisdictions, but are heavily influenced by the standards promulgated by the FDA and ICH.

All these stipulations are part of a quality standard called “Good Clinical Practice” (GCP) formally introduced by the ICH, but now generally extended to all regulatory systems. It consists of regulations (mandatory specifications, also called “Predicate Rules”) and guidelines (recommendations), ensuring that the RCT are scientifically valid and that the clinical properties of medicinal products are truthfully documented: The informed consent, protection of children, institutional review boards, disclosures of clinical investigators, trial protocols, investigator brochures, statistical principles etc.1,2

Computer Systems or software used to meet the requirements of these Predicate Rules need to show documented evidence that they “ensure accuracy, reliability, consistent intended performance, and the ability to discern invalid or altered records” 3, also known as Computer System Validation. In particular, the use of electronic records and electronic signatures in lieu of paper equivalents has been explicitly stipulated in 21 CFR 11, prescribing access control, record retention, audit trails, system checks and documentation of personnel roles and the meaning of signatures. There is extensive guidelines from the various authorities about the use and validation of these systems available 1-7.

## Regulations and guidelines for software used in observational studies

In contrast to this detailed system of regulations and compliance in RCT, the use of software for the collection and analysis of data in the course of observational studies is not at all covered by explicit regulations. The FDA only recently issued a “Guidance for Industry and FDA Staff- Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data” 8. Most of the principles laid out in this guideline are merely recommending to “consider” various factors potentially influencing the outcome of a study, and “encourage” to document these considerations and the choices made.

## Differences between RCTs and observational studies

There are several reasons for the observed discrepancies in regulations between RCTs and observational studies:

RCTs have been the main workhorse for testing medicinal products in humans, and their methodology has been refined over many decades, resulting in a large body of know-how with sufficient time to develop the regulatory environment. Observational studies on the other hand are a relatively new phenomenon, having only been possible at scale since the 1990s as they require large collections of data allowing sufficient sampling of the intervention and outcome.

In contrast to RCTs, these databases are not collected by the sponsors of the trials. Hence none of the regulations about collection of the data (many of the Predicate Rules) apply. Instead, they are collected by entirely different organizations for entirely different purposes: medical payers processing reimbursement claims or providers for managing patient care (Electronic Health Records, EHR). Observational studies are therefore “secondary use” of these data. The primary use is not regulated by the authorities, and the requirements for quality of the employed systems are entirely different.

In fact, observational data are known to almost always contain a small amount of inaccuracies, a few of which can easily be quantified through implausibilities, such as interventions on male reproductive organs in female patients, interventions on female reproductive organs in male populations and medical events after death, to name a few. In addition, records are often actively manipulated (removal of data, random date shifting) prior to secondary use for purposes of data protection, especially to prevent patient identification as required by the HIPAA regulations in the US 10 or equivalent stipulations in other countries.

Despite all these challenges, observational studies still fulfill an important role in the generation of evidence about medical interventions, for the reasons described above. The good news is that methods have been developed that can overcome many of the above shortcomings of the data through statistical means. The goal of OHDSI is the creation of a systematic, reproducible and standardized approach to conducting observational studies using such methods. Where the employed methods in conjunction with the imperfect data cannot guarantee the correct result, their performance should be known and the probability of producing the correct result quantified.

## Requirements for the OHDSI Methods Library

Given that there currently are only few regulations and guidelines specific to software used for observational research, we will attempt to infer them from other sources. Although the act of “validation” is interpreted in different ways in different fields, the FDA clearly defines the term as: “Establishing documented evidence which provides a high degree of assurance that a specific process will consistently produce a product meeting its predetermined specifications and quality attributes.” 11 For the OHDSI Methods Library, we interpret the *process* to mean the observational study, and the *product* to mean the estimates produced by the study, such as the effect size estimate in population-level estimation, and the predicted probability in patient-level prediction. One important quality attribute is calibration: we would like confidence intervals of effect size estimates to contain the true effect size approximately 95% of times, and we would like predicted probabilities to match observed frequencies.

The FDA explains 5 that validation encompasses the overall process and is designed to assure quality and consistency for a process throughout its lifecycle. In contrast, verification is an activity performed during and/or between phases of the overall lifecycle. Software testing is one form of verification. Accordingly, we identify characteristics of the OHDSI Methods Library development process that contribute to validity (section 5), and software testing performed on the OHDSI Methods Library (section 6).

# Software Development Process

The OHDSI Methods Library is developed by the OHDSI community. Proposed changes to the Library are discussed in two venues: The GitHub issue trackers (see Appendix A), and the OHDSI Forums 12. Both are open to the public. Any member of the community can contribute software code to the Library, however, final approval of any changes incorporated in the released versions of the software is performed by the OHDSI Population-Level Estimation Workgroup and OHDSI Patient-Level Prediction Workgroup leadership only.

Users can install the Methods Library in R directly from the master branches in the GitHub repositories, or through a system known as ‘drat’ that is always up-to-date with the master branches 14. A number of the Methods Library packages are available through R’s Comprehensive R Archive Network (CRAN), and this number is expected to increase over time.

Reasonable software development and testing methodologies are employed by OHDSI to maximize the accuracy, reliability and consistency of the Methods Library performance. Importantly, as the Methods Library is released under the terms of the Apache License V2, all source code underlying the Methods Library, whether it be in R, C++, SQL, or Java is available for peer review by all members of the OHDSI community, and the public in general. Thus, all the functionality embodied within Methods Library is subject to continuous critique and improvement relative to its accuracy, reliability and consistency.

## Source Code Management

All of the Methods Library’s source code is managed in the source code version control system ‘git’ publicly assessible via GitHub. The OHDSI Methods Library repositories are access controlled. Anyone in the world can view the source code, and any member of the OHDSI community can submit changes through so-called pull requests. Only the OHDSI Population-Level Estimation Workgroup and Patient-Level Prediction Workgroup leadership can approve such request, make changes to the master branches, and release new versions. Continuous logs of code changes are maintained within the GitHub repositories and reflect all aspects of changes in code and documentation. These commit logs are available for public review (see Appendix A).

New versions are released by the OHDSI Population-Level Estimation Workgroup and Patient-Level Prediction Workgroup leadership as needed. A new release starts by pushing changes to a master branch with a package version number (as defined in the DESCRIPTION file inside the package) that is greater than the version number of the previous release. This automatically triggers checking and testing of the package. If all tests are passed, the new version is automatically tagged in the version control system and the package is automatically uploaded to the OHDSI drat repository 14. New versions are numbered using three-component version number:

* New micro versions (e.g. from 4.3.2 to 4.3.3) indicate bug fixes only. No new functionality, and forward and backward compatibility are guaranteed
* New minor versions (e.g. from 4.3.3 to 4.4.0) indicate added functionality. Only backward compatibility is guaranteed
* New major versions (e.g. from 4.4.0 to 5.0.0) indicate major revisions. No guarantees are made in terms of compatibility

## Documentation

All packages in the Methods Library are documented through R’s internal documentation framework. Each package has a package manual that describes every function available in the package. To promote alignment between the function documentation and the function implementation, the roxygen2 software 13 is used to combine a function’s documentation and source code in a single file. The package manual is available on demand through R’s command line interface, or as a PDF in the package repositories (see Appendix A). In addition, many packages also have vignettes that highlight specific use cases of a package (see Appendix A).

All Method Library source code is available to end users. Feedback from the community is facilitated using GitHub’s issue tracking system (see Appendix A) and the OHDSI Forums 12.

## Availability of Current and Historical Archive Versions

Current and historical versions of the Methods Library packages are available in two locations: First, the GitHub version control system contains the full development history of each package (see Appendix A), and the state of a package at each point in time can be reconstructed and retrieved. Most importantly, each released version is tagged in GitHub. Second, the released R source packages are stored in the OHDSI GitHub drat repository 14.

## Maintenance, Support and Retirement

Each current version of the Methods Library is actively supported by OHDSI with respect to bug reporting, fixes and patches. Issues can be reported through GitHub’s issue tracking system, and through the OHDSI forums. Each package has a package manual, and zero, one or several vignettes. Online video tutorials are available 14, and in-person tutorials are provided from time to time.

## Qualified Personnel

As noted in Section 6.1, members of OHDSI community represent multiple statistical disciplines and are based at academic, not-for-profit and industry-affiliated institutions on multiple continents.

All leaders of the OHDSI Population-Level Estimation Workgroup and OHDSI Patient-Level Prediction Workgroup hold PhDs from accredited academic institutions and have published extensively in peer reviewed journals.

## Physical and Logical Security

The OHDSI Methods Library is hosted on the GitHub system. GitHub’s security measures are described elsewhere 16. Usernames and passwords are required by all members of the OHDSI community contribute modifications to the Methods Library, and only the Population-Level Estimation Workgroup and Patient-Level Prediction Workgroup leadership can makes changes to the master branches. User accounts are limited in access based upon standard security policies and functional requirements.

## Disaster Recovery

The OHDSI Methods Library is hosted on the GitHub system. GitHub’s disaster recovery facilities are described elsewhere 16.

# Testing

We distinguish between two types of tests performed on the Methods Library: Tests for individual functions in the packages (so-called ‘unit tests’), and tests to determine whether analyses implemented using the Methods Library produce reliable and accurate results (we will call this ‘method tests’).

## Unit test

A large set of automated validation tests is maintained and upgraded by OHDSI to enable the testing of source code against known data and known results. Each test begins with specifying some simple input data, then executes a function in one of the packages on this input, and evaluates whether the output is exactly what would be expected. For simple functions, the expected result is often obvious (for example when performing propensity score matching on example data containing only a few subjects), for more complicated functions the expected result may be generated using combinations of other functions available in R (for example, Cyclops, our large-scale regression engine, is tested amongst others by comparing results on simple problems with other regression routines in R). We aim for these tests in total to cover 100% of the lines of executable source code.

Appendix A lists the locations of the tests in each package. These tests are automatically performed when changes are made to a package (specifically, when changes are pushed to the package repository). Any errors noted during testing automatically trigger emails to the leadership of the Workgroups, and must be resolved prior to release of a new version of a package. The results of the unit tests can be found in the locations specified in Appendix A.

The source code and expected results for these tests are available for review and use in other applications as may be appropriate. These tests are also available to end users and/or system administrators and can be run as part of their installation process to provide further documentation and objective evidence as to the accuracy, reliability and consistency of their installation of the Methods Library.

## Method tests for population-level effect estimation

In addition, for the task of population-level effect estimation the Method Library is validated in its ability to produce effect estimates in line with the truth. We distinguish two types of tasks to evaluate: (1) estimation of the average effect of an exposure on an outcome relative to no exposure (*effect estimation*), and (2) estimation of the average effect of an exposure on an outcome relative to another exposure (*comparative effect estimation*). To evaluate the methods on these tasks, a gold standard was constructed consisting of 800 entries, with each item specifying a target exposure, comparator exposure, outcome, nesting cohort, and known true effect size. An example entry: target = Diclofenac, comparator = Celecoxib, outcome = Lyme disease, nesting cohort = Arthralgia, known true effect size (relative risk) = 1. Each entry can be used to evaluate for both tasks, since the true effect size holds both when comparing the target exposure to no exposure as well as when comparing the target exposure to the comparator exposure. The nesting cohort identifies a more homogeneous subgroup of a population, and can be used to evaluate methods such as the nested case-control design.

A subset of 200 entries are negative controls, where the relative risk is believed to be 1. The remaining 600 entries are positive controls, which were automatically derived from the 200 negative controls by adding synthetic additional outcomes during the target exposure until a desired incidence rate ratio was achieved between before and after injection of the synthetic outcomes 18. The target incidence rate ratios were 1.25, 2, and 4. To preserve (measured) confounding, predictive models were fitted for each outcome during target exposure and used to generate probabilities from which the synthetic outcomes were sampled.

After applying the methods on the gold standard the following metrics were computed:

* Area under the received operating curve (AUC), when comparing positive controls to negative controls
* Coverage of the 95% confidence interval of the effect estimates
* Mean precision of the effect estimates
* Mean squared error (MSE) of the effect estimates
* Type I and type II error in rejecting the null hypothesis regarding effect estimates

The full gold standard, as well as the computation of the metrics, can be found in the MethodEvaluation package which is part of the Methods Library. Results generated using a set of different analysis settings on the Truven MarketScan MDCD database, a large US insurance claims database, are included in Appendix B. Note that some methods inherently have poor performance that is not due to how they are implemented. For example, the case-control design in general does not adequately control for confounding, and therefore shows high type I error.

# Conclusions

The purpose of this report is to document evidence to provide a high degree of assurance that the Methods Library can be used in observational studies to consistently produce reliable and accurate estimates. Both through adoption of best software development practices during the software lifecycle, as well as continuous extensive testing of individual components of the software and the start-to-finish application of the methods library on a gold standard aim to ensure the validity of the Methods Library. However, use of the Methods Library does not guarantee validity of a study, since validity depends on many other components outside of the Methods Library as well, including appropriate study design, exposure and outcome definitions, and data quality. It is important to note that there is a significant obligation on the part of the end-user’s organization to define, create, implement and enforce the Method Library installation, validation and utilization related Standard Operating Procedures (SOPs) within the end-user’s environment. These SOPs should define appropriate and reasonable quality control processes to manage end-user related risk within the applicable operating framework. The details and content of any such SOPs are beyond the scope of this document.

# References

1. ICH Efficacy Guidelines: <http://www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html>
2. FDA Regulations Relating to Good Clinical Practice and Clinical Trials: <https://www.fda.gov/ScienceResearch/SpecialTopics/RunningClinicalTrials/ucm114928.htm>
3. Code of Federal Regulations Chapter 21, part 11: <https://www.ecfr.gov/cgi-bin/text-idx?node=pt21.1.11>
4. FDA Guidance for Industry - Computerized Systems Used in Clinical Investigations: <https://www.fda.gov/iceci/enforcementactions/bioresearchmonitoring/ucm135196.htm>
5. FDA General Principles of Software Validation: <https://www.fda.gov/MedicalDevices/ucm085281.htm>
6. ICH E9 - Statistical Principles for Clinical Trials: <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf>
7. Guidance for Industry and FDA Staff - Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials: <https://www.fda.gov/MedicalDevices/ucm071072.htm>
8. Guidance for Industry and FDA Staff- Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data: <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM243537.pdf>
9. The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM): <https://github.com/OHDSI/CommonDataModel>
10. The HIPAA Privacy Rule: <https://www.hhs.gov/hipaa/for-professionals/privacy/index.html>
11. FDA Glossary of Computer System Software Development Terminology: <https://www.fda.gov/iceci/inspections/inspectionguides/ucm074875.htm>
12. The OHDSI Forums: <http://forums.ohdsi.org/>
13. The roxygen2 package: <https://cran.r-project.org/web/packages/roxygen2/vignettes/roxygen2.html>
14. The OHDSI drat repository: <https://github.com/OHDSI/drat>
15. OHDSI Video Tutorials: <http://www.ohdsi.org/web/wiki/doku.php?id=videos:overview>
16. GitHub Security: <https://help.github.com/articles/github-security/>
17. Method Evaluation results: <http://data.ohdsi.org/MethodEvalViewer/>
18. Voss EA, Boyce RD, Ryan PB, van der Lei J, Rijnbeek PR, Schuemie MJ. Accuracy of an automated knowledge base for identifying drug adverse reactions. Journal of biomedical informatics. 2017;66:72-81.
19. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. Proceedings of the National Academy of Sciences. 2018.

# 

# 

# Appendix A: Method Library packages

## CaseControl

**Description:** CaseControl is an R package for performing (nested) matched case-control analyses in an observational database in the OMOP Common Data Model.

**Source code repository:**<https://github.com/OHDSI/CaseControl>

**Manual:**<https://raw.githubusercontent.com/OHDSI/CaseControl/master/extras/CaseControl.pdf>

**Vignette:** 'Multiple analyses using CaseControl' (<https://raw.githubusercontent.com/OHDSI/CaseControl/master/inst/doc/MultipleAnalyses.pdf>)

**Vignette:** 'Single studies using CaseControl' (<https://raw.githubusercontent.com/OHDSI/CaseControl/master/inst/doc/SingleStudies.pdf>)

**Issue tracker:**<https://github.com/OHDSI/CaseControl/issues>

**Commits log:**<https://github.com/OHDSI/CaseControl/commits/master>

**Unit tests:**<https://github.com/OHDSI/CaseControl/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/CaseControl>

## CaseCrossover

**Description:** An R package for performing case-crossover and case-time-control analyses in an observational database in the OMOP Common Data Model.

**Source code repository:**<https://github.com/OHDSI/CaseCrossover>

**Manual:**<https://raw.githubusercontent.com/OHDSI/CaseCrossover/master/extras/CaseCrossover.pdf>

**Vignette:** 'Multiple analyses using CaseCrossover' (<https://raw.githubusercontent.com/OHDSI/CaseCrossover/master/inst/doc/MultipleAnalyses.pdf>)

**Vignette:** 'Single studies using CaseCrossover' (<https://raw.githubusercontent.com/OHDSI/CaseCrossover/master/inst/doc/SingleStudies.pdf>)

**Issue tracker:**<https://github.com/OHDSI/CaseCrossover/issues>

**Commits log:**<https://github.com/OHDSI/CaseCrossover/commits/master>

**Unit tests:**<https://github.com/OHDSI/CaseCrossover/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/CaseCrossover>

## CohortMethod

**Description:** CohortMethod is an R package for performing new-user cohort studies in an observational database in the OMOP Common Data Model. It extracts the necessary data from a database in OMOP Common Data Model format, and uses a large set of covariates for both the propensity and outcome model, including for example all drugs, diagnoses, procedures, as well as age, comorbidity indexes, etc. Large scale regularized regression is used to fit the propensity and outcome models. Functions are included for trimming, stratifying and matching on propensity scores, as well as diagnostic functions, such as propensity score distribution plots and plots showing covariate balance before and after matching and/or trimming. Supported outcome models are (conditional) logistic regression, (conditional) Poisson regression, and (stratified) Cox regression.

**Source code repository:**<https://github.com/OHDSI/CohortMethod>

**Manual:**<https://raw.githubusercontent.com/OHDSI/CohortMethod/master/extras/CohortMethod.pdf>

**Vignette:** 'Multiple analyses using CohortMethod' (<https://raw.githubusercontent.com/OHDSI/CohortMethod/master/inst/doc/MultipleAnalyses.pdf>)

**Vignette:** 'Single studies using CohortMethod' (<https://raw.githubusercontent.com/OHDSI/CohortMethod/master/inst/doc/SingleStudies.pdf>)

**Issue tracker:**<https://github.com/OHDSI/CohortMethod/issues>

**Commits log:**<https://github.com/OHDSI/CohortMethod/commits/master>

**Unit tests:**<https://github.com/OHDSI/CohortMethod/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/CohortMethod>

## Cyclops

**Description:** This model fitting tool incorporates cyclic coordinate descent and majorization-minimization approaches to fit a variety of regression models found in large-scale observational healthcare data. Implementations focus on computational optimization and fine-scale parallelization to yield efficient inference in massive datasets.

**Source code repository:**<https://github.com/OHDSI/Cyclops>

**Manual:**<https://raw.githubusercontent.com/OHDSI/Cyclops/master/extras/Cyclops.pdf>

**Issue tracker:**<https://github.com/OHDSI/Cyclops/issues>

**Commits log:**<https://github.com/OHDSI/Cyclops/commits/master>

**Unit tests:**<https://github.com/OHDSI/Cyclops/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/Cyclops>

## DatabaseConnector

**Description:** Package for connecting to various DBMSs. Also includes support for fetching data as ffdf objects.

**Source code repository:**<https://github.com/OHDSI/DatabaseConnector>

**Manual:**<https://raw.githubusercontent.com/OHDSI/DatabaseConnector/master/extras/DatabaseConnector.pdf>

**Issue tracker:**<https://github.com/OHDSI/DatabaseConnector/issues>

**Commits log:**<https://github.com/OHDSI/DatabaseConnector/commits/master>

**Unit tests:**<https://github.com/OHDSI/DatabaseConnector/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/DatabaseConnector>

## EmpiricalCalibration

**Description:** Routines for performing empirical calibration of observational study estimates. By using a set of negative control hypotheses we can estimate the empirical null distribution of a particular observational study setup. This empirical null distribution can be used to compute a calibrated p-value, which reflects the probability of observing an estimated effect size when the null hypothesis is true taking both random and systematic error into account.

**Source code repository:**<https://github.com/OHDSI/EmpiricalCalibration>

**Manual:**<https://raw.githubusercontent.com/OHDSI/EmpiricalCalibration/master/extras/EmpiricalCalibration.pdf>

**Vignette:** 'Empirical calibration of confidence intervals' (<https://raw.githubusercontent.com/OHDSI/EmpiricalCalibration/master/inst/doc/EmpiricalCiCalibrationVignette.pdf>)

**Vignette:** 'Empirical calibration of p-values' (<https://raw.githubusercontent.com/OHDSI/EmpiricalCalibration/master/inst/doc/EmpiricalPCalibrationVignette.pdf>)

**Issue tracker:**<https://github.com/OHDSI/EmpiricalCalibration/issues>

**Commits log:**<https://github.com/OHDSI/EmpiricalCalibration/commits/master>

**Unit tests:**<https://github.com/OHDSI/EmpiricalCalibration/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/EmpiricalCalibration>

## EvidenceSynthesis

**Description:** Routines for combining evidence and diagnostics across multiple sources, such as multiple data sites in a distributed study. This includes functions for performing meta-analysis and forest plots.

**Source code repository:**<https://github.com/OHDSI/EvidenceSynthesis>

**Manual:**<https://raw.githubusercontent.com/OHDSI/EvidenceSynthesis/master/extras/EvidenceSynthesis.pdf>

**Issue tracker:**<https://github.com/OHDSI/EvidenceSynthesis/issues>

**Commits log:**<https://github.com/OHDSI/EvidenceSynthesis/commits/master>

**Unit tests:**<https://github.com/OHDSI/EvidenceSynthesis/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/EvidenceSynthesis>

## FeatureExtraction

**Description:** An R package for generating features (covariates) for a cohort using data in the Common Data Model.

**Source code repository:**<https://github.com/OHDSI/FeatureExtraction>

**Manual:**<https://raw.githubusercontent.com/OHDSI/FeatureExtraction/master/extras/FeatureExtraction.pdf>

**Vignette:** 'Creating covariates using cohort attributes' (<https://raw.githubusercontent.com/OHDSI/FeatureExtraction/master/inst/doc/CreatingCovariatesUsingCohortAttributes.pdf>)

**Vignette:** 'Creating custom covariate builders' (<https://raw.githubusercontent.com/OHDSI/FeatureExtraction/master/inst/doc/CreatingCustomCovariateBuilders.pdf>)

**Vignette:** 'Using FeatureExtraction' (<https://raw.githubusercontent.com/OHDSI/FeatureExtraction/master/inst/doc/UsingFeatureExtraction.pdf>)

**Issue tracker:**<https://github.com/OHDSI/FeatureExtraction/issues>

**Commits log:**<https://github.com/OHDSI/FeatureExtraction/commits/master>

**Unit tests:**<https://github.com/OHDSI/FeatureExtraction/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/FeatureExtraction>

## MethodEvaluation

**Description:** This package contains resources for the evaluation of the performance of methods that aim to estimate the magnitude (relative risk) of the effect of a drug on an outcome. These resources include reference sets for evaluating methods on real data, as well as functions for inserting simulated effects in real data based on negative control drug-outcome pairs. Further included are functions for the computation of the minimum detectable relative risks and functions for computing performance statistics such as predictive accuracy, error and bias.

**Source code repository:**<https://github.com/OHDSI/MethodEvaluation>

**Manual:**<https://raw.githubusercontent.com/OHDSI/MethodEvaluation/master/extras/MethodEvaluation.pdf>

**Issue tracker:**<https://github.com/OHDSI/MethodEvaluation/issues>

**Commits log:**<https://github.com/OHDSI/MethodEvaluation/commits/master>

**Unit tests:**<https://github.com/OHDSI/MethodEvaluation/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/MethodEvaluation>

## OhdsiRTools

**Description:** Format and check syntax of R code and packages following the OHDSI R style guidelines. Support for parallel computation.

**Source code repository:**<https://github.com/OHDSI/OhdsiRTools>

**Manual:**<https://raw.githubusercontent.com/OHDSI/OhdsiRTools/master/extras/OhdsiRTools.pdf>

**Issue tracker:**<https://github.com/OHDSI/OhdsiRTools/issues>

**Commits log:**<https://github.com/OHDSI/OhdsiRTools/commits/master>

**Unit tests:**<https://github.com/OHDSI/OhdsiRTools/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/OhdsiRTools>

## PatientLevelPrediction

**Description:** A package for creating patient level prediction models. Given a cohort of interest and an outcome of interest, the package can use data in the OMOP Common Data Model to build a large set of features. These features can then be assessed to fit a predictive model using a number of machine learning algorithms. Several performance measures are implemented for model evaluation.

**Source code repository:**<https://github.com/OHDSI/PatientLevelPrediction>

**Manual:**<https://raw.githubusercontent.com/OHDSI/PatientLevelPrediction/master/extras/PatientLevelPrediction.pdf>

**Issue tracker:**<https://github.com/OHDSI/PatientLevelPrediction/issues>

**Commits log:**<https://github.com/OHDSI/PatientLevelPrediction/commits/master>

**Unit tests:**<https://github.com/OHDSI/PatientLevelPrediction/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/PatientLevelPrediction>

## SelfControlledCaseSeries

**Description:** SelfControlledCaseSeries is an R package for performing self- controlled case series (SCCS) analyses in an observational database in the OMOP Common Data Model.

**Source code repository:**<https://github.com/OHDSI/SelfControlledCaseSeries>

**Manual:**<https://raw.githubusercontent.com/OHDSI/SelfControlledCaseSeries/master/extras/SelfControlledCaseSeries.pdf>

**Vignette:** 'Multiple analyses using SelfControlledCaseSeries' (<https://raw.githubusercontent.com/OHDSI/SelfControlledCaseSeries/master/inst/doc/MultipleAnalyses.pdf>)

**Vignette:** 'Single studies using SelfControlledCaseSerie' (<https://raw.githubusercontent.com/OHDSI/SelfControlledCaseSeries/master/inst/doc/SingleStudies.pdf>)

**Issue tracker:**<https://github.com/OHDSI/SelfControlledCaseSeries/issues>

**Commits log:**<https://github.com/OHDSI/SelfControlledCaseSeries/commits/master>

**Unit tests:**<https://github.com/OHDSI/SelfControlledCaseSeries/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/SelfControlledCaseSeries>

## SelfControlledCohort

**Description:** This package provides a method to estimate risk by comparing time exposed with time unexposed among the exposed cohort.

**Source code repository:**<https://github.com/OHDSI/SelfControlledCohort>

**Manual:**<https://raw.githubusercontent.com/OHDSI/SelfControlledCohort/master/extras/SelfControlledCohort.pdf>

**Issue tracker:**<https://github.com/OHDSI/SelfControlledCohort/issues>

**Commits log:**<https://github.com/OHDSI/SelfControlledCohort/commits/master>

**Unit tests:**<https://github.com/OHDSI/SelfControlledCohort/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/SelfControlledCohort>

## SqlRender

**Description:** A rendering tool for parameterized SQL that also translates into different SQL dialects. These dialects include Sql Server, Oracle, PostgreSql, Amazon RedShift, Impala, IBM Netezza, Google BigQuery, and Microsoft PDW.

**Source code repository:**<https://github.com/OHDSI/SqlRender>

**Manual:**<https://raw.githubusercontent.com/OHDSI/SqlRender/master/extras/SqlRender.pdf>

**Vignette:** 'Using SqlRender' (<https://raw.githubusercontent.com/OHDSI/SqlRender/master/inst/doc/UsingSqlRender.pdf>)

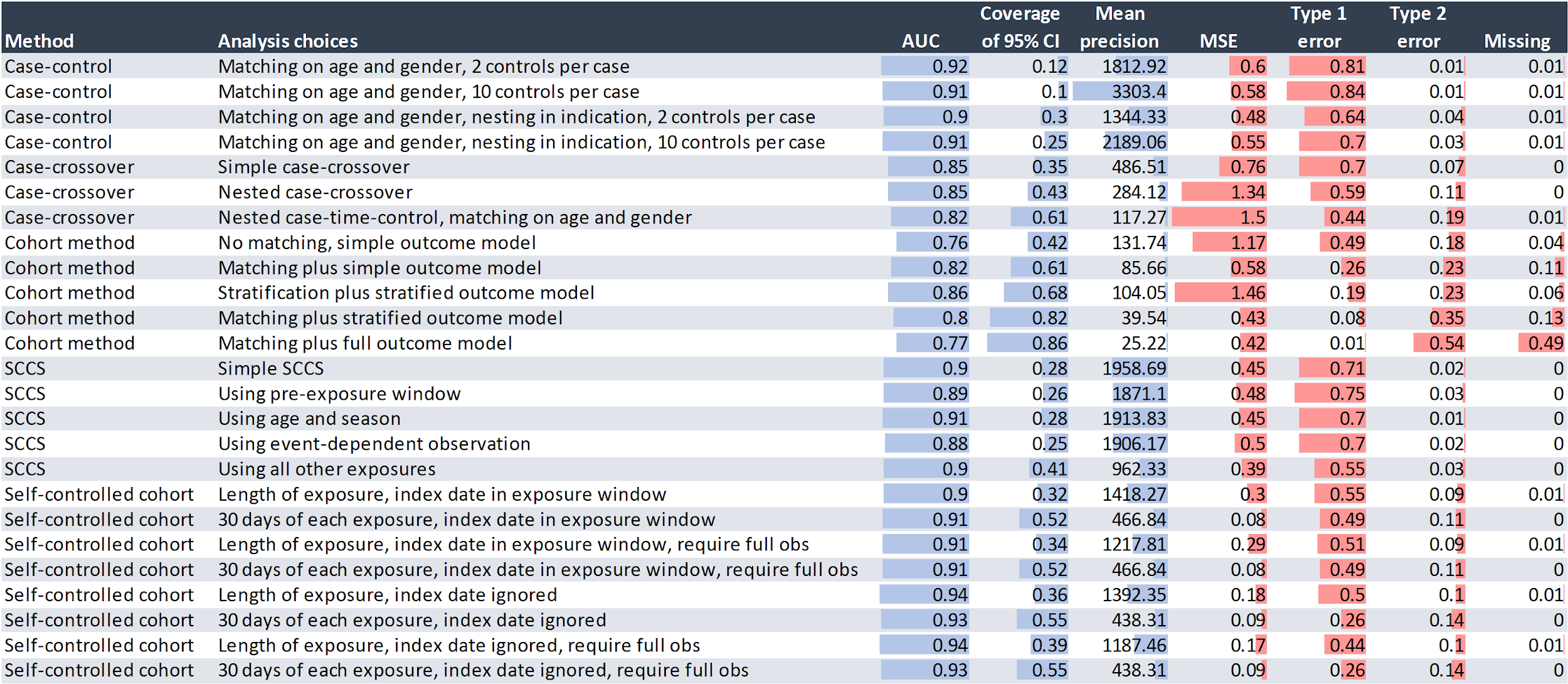
**Issue tracker:**<https://github.com/OHDSI/SqlRender/issues>

**Commits log:**<https://github.com/OHDSI/SqlRender/commits/master>

**Unit tests:**<https://github.com/OHDSI/SqlRender/tree/master/tests/testthat>

**Unit test results:**<https://travis-ci.org/OHDSI/SqlRender>

# Appendix B: Results of the method evaluation



**Table 1.** Performance metrics of various study designs implemented using the OHDSI Methods Library. Performance was measured using the OHDSI Methods Benchmark as contained in the MethodEvaluation package and was generated using the Truven MDCD database. AUC: Area under the ROC curve for classifying all positive controls vs. all negative controls, Coverage: Coverage of the 95% confidence interval, Mean precision: Precision = 1/variance; higher precision means narrower confidence intervals, MSE: Mean squared error between effect size (point) estimate and the true effect size, Type 1 error: For all negative controls, how often was the null rejected (at alpha = 0.05), Type 2 error: For all positive controls, how often was the null not rejected (at alpha = 0.05), Missing: For how many of the controls was the method unable to produce an estimate