

# MINI-SENTINEL METHODS

## PROPENSITY SCORE MATCHING TOOL ENHANCEMENTS

### REPORT

**Prepared by:** Joshua J Gagne, PharmD, ScD,<sup>1</sup> Martin Kulldorff, PhD,<sup>4</sup> Shirley V. Wang, PhD, ScM,<sup>1</sup> Ivair Silva, PhD,<sup>6</sup> Bruce Fireman, MS,<sup>3</sup> Wesley Eddings, PhD,<sup>1</sup> Kunthel By, PhD,<sup>2</sup> Rita Ouellet-Hellstrom, PhD, MPH,<sup>5</sup> Simone Pinheiro, ScD, MSc,<sup>5</sup> Jennifer R. Popovic, DVM, MA,<sup>4</sup> Marsha Reichman, PhD,<sup>5</sup> Catherine Panozzo, PhD, MPH,<sup>4</sup> Azadeh Shoaibi, PhD, MS, MHS,<sup>7</sup> Sebastian Schneeweiss, MD, ScD<sup>1</sup>

**Author Affiliations:** 1. Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, Boston, MA 2. Office of Translational Sciences, Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD 3. Division of Research, Kaiser Permanente Northern California, Oakland, CA 4. Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA 5. Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD 6. Department of Statistics, Universidade Federal de Ouro Preto, Ouro Preto, Minas Gerais, Brazil 7. Office of Medical Policy, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD

**September 21, 2015**

Mini-Sentinel is a pilot project sponsored by the [U.S. Food and Drug Administration \(FDA\)](#) to inform and facilitate development of a fully operational active surveillance system, the Sentinel System, for monitoring the safety of FDA-regulated medical products. Mini-Sentinel is one piece of the [Sentinel Initiative](#), a multi-faceted effort by the FDA to develop a national electronic system that will complement existing methods of safety surveillance. Mini-Sentinel Collaborators include Data and Academic Partners that provide access to health care data and ongoing scientific, technical, methodological, and organizational expertise. The Mini-Sentinel Coordinating Center is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223200910006I.

## Mini-Sentinel Methods

### Propensity Score Matching Tool Enhancements Report

#### Table of Contents

<b>I.</b>	<b>INTRODUCTION.....</b>	<b>- 1 -</b>
<b>II.</b>	<b>OBJECTIVES.....</b>	<b>- 2 -</b>
<b>III.</b>	<b>EXPAND SEQUENTIAL TESTING CAPABILITIES TO VARIABLE-RATIO MATCHED COHORT SETTING..</b>	<b>- 2 -</b>
<b>IV.</b>	<b>RISK SET-LEVEL DATA AGGREGATION PROCESS .....</b>	<b>- 3 -</b>
A.	RISK-SET-LEVEL DATA THAT EACH DATA PARTNER WILL TRANSMIT TO MSOC.....	- 4 -
B.	STEPS FOR ACCOMMODATING RISK SET-LEVEL DATA AGGREGATION .....	- 5 -
1.	<i>Build the new risk set-level dataset off of the individual level data .....</i>	<i>- 5 -</i>
2.	<i>Add case-centered logistic regression model to analysis code.....</i>	<i>- 6 -</i>
<b>V.</b>	<b>DEVELOP AND IMPLEMENT STRATEGY FOR PRODUCING DISTRIBUTED KAPLAN-MEIER PLOTS .....</b>	<b>- 6 -</b>
A.	MINIMUM DATA REQUIRED FROM EACH DATA PARTNER TO CONSTRUCT A KAPLAN-MEIER PLOT.....	- 7 -
<b>VI.</b>	<b>IDENTIFY AND PRIORITIZE ADDITIONAL POTENTIAL ENHANCEMENTS.....</b>	<b>- 8 -</b>
A.	ENHANCEMENTS THAT REQUIRE SCIENTIFIC LEADERSHIP.....	- 8 -
B.	MSOC ENHANCEMENTS.....	- 8 -
C.	DATA CHARACTERIZATION.....	- 10 -
D.	REPORTING .....	- 10 -
E.	PROCESS.....	- 10 -
<b>VII.</b>	<b>ACKNOWLEDGEMENTS.....</b>	<b>- 12 -</b>
<b>VIII.</b>	<b>REFERENCES.....</b>	<b>- 13 -</b>

## I. INTRODUCTION

The Mini-Sentinel pilot recently developed and introduced a suite of Prospective Routine Observational Monitoring Program Tools (PROMPTs) to enable FDA to perform sequential surveillance of health outcomes related to new products as patient experience with these products accrues in the Mini-Sentinel Distributed Database.<sup>1</sup> The PROMPTs include four modules – three that use a cohort-based design and one that uses a self-controlled design.<sup>2-5</sup> FDA is currently using one of the cohort-based approaches – the Propensity Score Matching (PSM) module – in ongoing assessments of newly marketed medications.<sup>6,7</sup>

The PSM module comprises a suite of flexible, scalable, modular SAS macros that perform effect estimation in a distributed data setting based on propensity-score matched sequential new user cohorts.<sup>8</sup> The PSM module was originally built on top of Mini-Sentinel’s Modular Program 3 (MP3), which was a program that created cohorts of new users for one or more drug/procedures of interest and indicators of incident events.<sup>9</sup> The stand-alone MP3 has since been retired within Mini-Sentinel and the PSM module has now been integrated into the newer Cohort Identification and Descriptive Analysis (CIDA) tool.<sup>10</sup> The adjustment component of the PSM module takes the cohort of new users identified by CIDA and creates indicators for predefined covariate conditions, procedures, and concomitant drug exposures (defined through diagnostic codes, procedure codes and prescription medication claims) and calculates a combined Charlson-Elixhauser comorbidity score.<sup>11</sup> The program has the option for the use of the high-dimensional propensity score (hdPS) algorithm to identify empirical confounders as well.<sup>12</sup> The matching component of the module performs both fixed and variable-ratio propensity score matching using a nearest-neighbor matching algorithm with any of three user-specified maximum matching calipers (i.e., 0.01, 0.025, and 0.05 on the propensity score scale).

The PSM tool produces datasets that remain behind Data Partner firewalls and others that are returned to the Mini-Sentinel Operations Center (MSOC). The de-identified datasets that each Data Partner returns to the MSOC are used in a site-stratified Cox proportional hazards model to estimate an overall hazard ratio across Data Partners. This de-identified dataset includes an indicator of exposure status, an indicator of whether patients experienced the outcome of interest, a variable indicating the number of days of follow-up, and a propensity score value, which summarizes into a single scalar variable all

demographic, clinical, and health service utilization variables necessary for confounding adjustment. If requested, subgroup indicators for selected covariates can be returned to the MSOC. No patient identifiers are returned in this dataset.

The PSM tool also automatically returns a number of output files, including Tables 1 from both matched and unmatched new user cohorts from each Data Partner and propensity score distributions before and after matching from each Data Partner. The module also produces as an output file the aggregate exposure and outcome data needed for the maximum sequential probability ratio test (maxSPRT), which is a likelihood-ratio test for binomial data. This file is fed into the 'R sequential' software package developed as part of the self-controlled PROMPT to perform sequential testing in the matched cohort.

## **II. OBJECTIVES**

This Workgroup addressed four specific aims related to expanding the capabilities of the PROMPT PSM module:

1. Expand sequential testing capabilities to variable-ratio matched cohort setting (Section III)
2. Incorporate risk-set level data aggregation process (Section IV)
3. Develop and implement strategy for producing distributed Kaplan-Meier plots (Section V)
4. Identify and prioritize additional potential enhancements (Section VI)

## **III. EXPAND SEQUENTIAL TESTING CAPABILITIES TO VARIABLE-RATIO MATCHED COHORT SETTING**

The PSM tool enables estimation of multiple parameters, including both hazard ratios and incidence rate differences based on either fixed or variable patient follow-up periods. The tool uses the same maxSPRT sequential testing procedure and the same R Sequential package that is used in the PROMPT Self-Controlled tool.<sup>13</sup> The matching ratio must be fixed to apply the original R sequential functions for binomial data; that is, the number of patients exposed to the comparator product who are matched to patients exposed to the product of interest must be constant (e.g., 1:1 or 1:2). While fixed 1:1 matching (which is the default in the current module) ensures the best balance in covariates between treatment

groups, variable ratio matching (i.e., matching more than one member in one treatment group to each member of the other group) can increase statistical power.

The Workgroup developed extensions to the R Sequential binomial functions to accommodate variable ratio matching. The extensions have been implemented in the free 'R Sequential' package (<http://cran.r-project.org/web/packages/Sequential/index.html>). Currently, the new functions are available for FDA and MSOC and they will be part of the next official release of the R Sequential package, which is scheduled for early June 2015.

#### IV. RISK SET-LEVEL DATA AGGREGATION PROCESS

The PROMPT PSM tool generates de-identified, individual-level datasets (including only exposure status, outcome status, propensity score, and indicators for subgroup variables) at each Data Partner. These files are then transmitted to MSOC where they are appended across Data Partners. A site-stratified Cox proportional hazards regression model is used to analyze the appended data sets. A legal expert reviewed a prototype of the dataset that is transferred and confirmed that it complies with HIPAA (Rassen JA et al. Evaluating strategies for data sharing and analyses in distributed data settings. Mini-Sentinel 2013. Available at: [http://www.mini-sentinel.org/methods/methods\\_development/details.aspx?ID=1041](http://www.mini-sentinel.org/methods/methods_development/details.aspx?ID=1041)). Despite this HIPAA-compliant data-transfer, one Data Partner indicated that it was not able to transmit data in this file format for a portion of their covered population, namely those for which they provided administrative services only (ASO). ASO arrangements are those in which an employer funds its own health insurance program but hires another organization to perform specific administrative services, such as claims processing.

To overcome this issue, the Workgroup implemented an alternative approach that rolls the individual-level data into risk-set-level data, before being transmitted. This case-centered logistic regression approach permits exactly the same propensity-score matched analysis with Cox proportional hazards regression, but achieves this without the need for individual-level data.<sup>14</sup> **Table 1** provides revised sample data for a hypothetical cohort that would be transmitted by each Data Partner to the MSOC using the risk set-level data aggregation process. MSOC confirmed with the respective Data Partner that

transmission of data at this level was permissible for the ASO population. Using this approach, data transfer for both ASO and non-ASO populations would be accomplished using risk set-level data.<sup>14</sup>

#### A. RISK-SET-LEVEL DATA THAT EACH DATA PARTNER WILL TRANSMIT TO MSOC

**Table 1** provides revised sample data for a hypothetical cohort that would be transmitted by each Data Partner to the MSOC using the risk set-level data aggregation process. Each row represents a risk set, which is defined by a propensity score-matched set in which one or more patients experienced the outcome of interest. For each risk set, only information about the number of days following the index date that the outcome occurred, the case’s exposure status, the probability of exposure among all patients in the risk set is needed. For example, in the first risk set, (Risk set ID = 1), the outcome occurred 73 days after treatment initiation (Follow-up days = 73) and the case was exposed to the drug of interest (Case exposure status = 1), and one in five patients in the risk set were exposed (Exposure probability in risk set = 0.2000). Note that non-informative risk sets (e.g., those in which no outcome occurs) are not transmitted.

Table 1. Example of Risk-Set-Level Data That Each Data Partner Transmits to MOSC						
Data Partner ID	Matching strategy	Risk set ID	Matched set ID	Case exposure status	Exposure probability in risk set*	Follow-up days
1	1	1	1	1	0.2000	73
1	1	2	2	0	0.5000	44
1	1	3	2	0	0.0000	163
1	1	4	3	0	0.3333	101
1	...	...	...	...	...	...
1	18	<i>j</i>	<i>n</i>	1	1.0000	88

Note that ‘exposure probability’ refers to the probability that patients were exposed to the drug of interest (e.g., new drug) versus the comparator.

Workgroup members worked closely with MSOC and the developers to implement this enhancement. MSOC confirmed with the respective Data Partner that transmission of data at this level was permissible for the ASO population. Using this approach, data transfer for both ASO and non-ASO populations would be accomplished using risk set-level data. Use of the risk-set level data approach therefore

enables the inclusion of a potentially larger sample but the de-identified individual-level data approach permits more analytic flexibility at the MSOC, such as the ability to apply different propensity score matching strategies.

## **B. STEPS FOR ACCOMMODATING RISK SET-LEVEL DATA AGGREGATION**

The steps below outline the changes made to the PSM tool to accommodate the risk set-level data aggregation enhancement. Programming code and technical specifications for the original PSM tool are available on the Mini-Sentinel webpage ([http://www.mini-sentinel.org/methods/methods\\_development/details.aspx?ID=1045](http://www.mini-sentinel.org/methods/methods_development/details.aspx?ID=1045)).

### **1. Build the new risk set-level dataset off of the individual level data**

This is done at the Data Partner site before any data transfer to the MSOC. The new code deletes uninformative matched sets (i.e., those sets in which no patient has the event) from the original individual-level data set that would have been transmitted if the risk set-level data approach were not used. For each matched set, the code determines whether any patient had an event and, if no patient in the set had an event, it deletes the whole set because it does not contribute to the partial likelihood of the Cox model. From matched sets with at least one event, the code creates risk-set level observations for each informative matched set by first identifying the first event in the matched set, as defined by the event corresponding to the smallest follow-up time. The number of days between the index date (i.e., drug initiation date) and the event date is added to the risk set-level data set (i.e., Follow-up days variable). The exposure status of the case is also added to the risk set-level data set (i.e., Case exposure status). Among all patients in the matched set with follow-up time greater than or equal to the follow-up time for the case (including the case itself; these patients form the risk set), the code calculates the probability of exposure by dividing the number of exposed people by the total number of people in the risk set to get the probability of exposure. This exposure probability is added to the risk set-level data set (i.e., Exposure probability in risk set). If the risk set contains more than one event, the code repeats this process of forming risk sets and determining the case exposure status, the follow-up days, and the exposure probability for each event. A separate observation is created for each case-defined risk set.

## 2. Add case-centered logistic regression model to analysis code

As described above, the analysis of the individual-level data sets is done at the MSOC using standardized code that appends the data sets from each Data Partner and fits a site-stratified Cox proportional hazard model. This functionality is retained in the enhanced code, but an additional model – the case-centered logistic regression model – is added to the standardized code to analyze the risk set-level data. This is accomplished by first creating a variable to convert the exposure probability in the risk set to log odds of exposure by taking the natural log of the odds of exposure, where the odds of exposure is exposure probability divided by one minus exposure probability. This variable is then used as an offset in the logistic regression model, in which the case’s exposure status in each risk set is the dependent variable.

The risk set level approach enables all of the diagnostics of the original PSM tool, including propensity score distribution figures and tables depicting distributions of baseline covariates between exposure groups. It is important to note that the risk set-level data aggregation enhancement is implemented as an alternative to use of the individual-level dataset, so that the tool retains the option of using either approach.

## V. DEVELOP AND IMPLEMENT STRATEGY FOR PRODUCING DISTRIBUTED KAPLAN-MEIER PLOTS

During the project period, the Workgroup identified that the addition of a Kaplan-Meier plot as a standard PSM tool output would be a high-priority enhancement. An important consequence of the incorporation of the risk set-level data aggregation process is that the risk set-level dataset that the Data Partners transmit to MSOC does not contain the necessary information for constructing Kaplan-Meier plots. The Workgroup considered several options for constructing Kaplan-Meier plots in the setting of distributed data analyses where individual-level data are not returned by Data Partners.

The Workgroup determined that an aggregate Kaplan-Meier plot could be constructed using minimal information (see **Table 2**). The only information required is the number of individuals in each exposure group who remain under observation at each follow-up day relative to the index, drug initiation date,



and the number of these individuals who experience the event of interest on each follow-up day. Numbers of patients and outcomes in each group on each follow-up day can then be summed across contributing Data Partners. This provides the necessary information to construct exactly the same Kaplan-Meier plot that can be constructed using individual-level data returned by each Data Partner.

**A. MINIMUM DATA REQUIRED FROM EACH DATA PARTNER TO CONSTRUCT A KAPLAN-MEIER PLOT**

<b>Table 2. Example of Minimum Data Required from Each Data Partner to Construct a Kaplan-Meier Plot</b>				
<b>Follow-up day</b>	<b>Exposed patients at risk</b>	<b>Exposed events</b>	<b>Unexposed patients at risk</b>	<b>Unexposed events</b>
1	10000	0	10000	0
2	10000	0	9999	1
3	9999	0	9997	1
4	9996	1	9993	0
5	9990	1	9989	1
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
365	5000	1	4500	1

MSOC confirmed with the relevant Data Partner that transmission of data at this level was permissible for the ASO population. The creation of this table is being included as part of the risk set-level data aggregation enhancement described above.

## **VI. IDENTIFY AND PRIORITIZE ADDITIONAL POTENTIAL ENHANCEMENTS**

The fourth objective of the Workgroup was to identify and prioritize future enhancements to propensity score-based tools, which are listed below. Note that the MSOC envisions combining the propensity score-based tools (e.g., PSM and IPTW modules) into a single tool that includes the functionality of both individual programs. The potential enhancements below are separated into those that require scientific leadership, those that would be the responsibility of the MSOC, those related to data characterization, and those related to the reporting tools. Note that all enhancements would be compatible with both individual-level and risk set-level data set approaches.

### **A. ENHANCEMENTS THAT REQUIRE SCIENTIFIC LEADERSHIP**

- Develop optional risk set level data return approach for propensity score-stratified analysis
  - Add parameter to enable requestor-specified number of strata (e.g., quintiles, deciles, vigintiles, centiles, etc).
- Add ability to perform PS trimming
- Develop additional approaches for incorporating propensity score into the analysis (inverse probability weighting and regression)
- For 1:1 matching, add the capability to stratify by follow-up time to explore time-varying hazards
- Add ability to define multiple comparison groups. Would allow n-wise (vs. pair wise) comparison. Would require changes to PSM calculation.

### **B. MSOC ENHANCEMENTS**

- Add parameters to enable inclusion of any variable and fixed ratio matching cap (e.g., 2, 3 for fixed or 5, 10 for variable)

- For 1:1 matching, add the capability to do unconditional analysis (worth noting that the unconditional analysis can only be done for fixed-ratio matching (e.g., fixed 1:1, 2:1, 3:1, etc). Conditional analysis is required for variable matching.
- Make specification of health service utilization covariates in the PS estimation models be requester-defined (i.e., no longer automatically include variables into the model)
- Add ability to include TIME and YEAR to Table 1 output
- Create Table 1 output for variable matching ratio option
- Add ability to adjust for additional covariates in the model used to estimate the HR and 95% CIs (for covariates that, based on standardized difference information, may not be adjusted for adequately by matching on PS). This is likely only possible for assessments that request patient-level information.
- Add more flexibility for defining covariates:
  - Add ability to define different look back windows for covariates, and optionally include the index date in the look back window. This requires CIDA tool modification and would benefit any future analyses that use the analytic dataset output by CIDA.
  - Add ability to select primary, secondary, unspecified positions for covariate
  - Add ability to specify location of codes (e.g., inpatient, outpatient, emergency department) for covariates
  - Enable more complex covariate definitions (e.g., requirement of 2 outpatient diagnosis codes; presence of a diagnosis code OR a drug dispensation)
- Enhance flexibility of subgroup analyses: currently, the program can accommodate subgroup analyses of any variable included in the PS model. For continuous variables (e.g., age), the same categories are used for both confounding adjustment in the PS model and for defining the subgroup categories. This modification would enable separate forms of the variables for inclusion in the PS model and for defining subgroups.

### C. DATA CHARACTERIZATION

- Create a wider Table 1 to profile baseline covariates over time to "diagnose" whether there may be differential censoring; the wider Table 1 can be created at a user-defined follow-up time point, such as median outcome event time
- Allow comparisons (details to be specified) between patients in time interval 1 in look 1 and patients in time interval 1 in look 2 (to observe potential data changes) to understand data changes across ETLs and impact on analyses
- Create Table 1 for each subgroup

### D. REPORTING

- Show number of matched controls as a function of propensity score (alternative approach is to define matching caliper as a proportion (e.g., 0.2) of the standard deviation of the logit of the propensity score)
- Output distribution of number of matches per patient for variable matching ratio requests
- Create forest plots (for site-specific effect estimates) and include in standard reporting
- Create histogram of number of controls in each matched set (To get some idea about the ratios in the variable matching)
- Add columns indicating informative events in the Cox (or case-centered logistic regression) model in addition to total analysis-eligible events in the cohorts (note, these are currently produced in a separate file that is used in the sequential R code)
- Remove Absolute Difference and Standardized Difference from baseline characteristic tables for the following variables: Patients, Events while on therapy, Mean person-days at risk
- Change the footnote ("standard deviation where no % follows the value") to "mean and standard deviation where no % follows the value" and add the superscript 1 to the "N" columns, too
- Expand number of decimal places to 2 for Absolute Difference and Standardized Difference

### E. PROCESS

- Move matching from Java to SAS

- Move hdPS from Java to SAS
- Eliminate need for Excel “Table 1 template”. MSOC should receive the information as a dataset and create a local aggregation and reporting program.

Workgroup leads will continue to work with MSOC to accommodate above mentioned enhancements in future versions of the PROMPT PSM tool.

## **VII. ACKNOWLEDGEMENTS**

We thank Sophia Axtman and Susan Farrow for their effective project management and support.

## VIII. REFERENCES

1. Prospective Routine Observational Monitoring Program Tools (PROMPT) Users' Guide. Mini-Sentinel 2014; available at [http://www.mini-sentinel.org/work\\_products/Statistical Methods/Mini-Sentinel PROMPT Users-Guide.pdf](http://www.mini-sentinel.org/work_products/Statistical Methods/Mini-Sentinel PROMPT Users-Guide.pdf).
2. Mini-Sentinel Prospective Routine Observational Monitoring Program Tool: Creation of the Analytic Dataset and Conducting Analysis for the Group Sequential Generalized Estimating Equation (GS GEE) Tool Technical Users' Guide version 1.0. Mini-Sentinel 2014; available at [http://www.mini-sentinel.org/work\\_products/Statistical Methods/Mini-Sentinel PROMPT Group-Sequential-Generalized-Estimating-Equation-Tool Users-Guide.pdf](http://www.mini-sentinel.org/work_products/Statistical Methods/Mini-Sentinel PROMPT Group-Sequential-Generalized-Estimating-Equation-Tool Users-Guide.pdf).
3. Cook A, Wellman R, Nelson J. Mini-Sentinel Prospective Routine Observational Monitoring Program Tool: Creation of the Analytic Dataset and Conducting Analysis for the Group Sequential Inverse Probability of Treatment Weighted (GS IPTW) Tool Technical Users' Guide version 1.0. Mini-Sentinel 2014; available at [http://www.mini-sentinel.org/work\\_products/Statistical Methods/Mini-Sentinel PROMPT Group-Sequential-Inverse-Probability-of-Treatment-Weighted-Regression-Tool Users-Guide.pdf](http://www.mini-sentinel.org/work_products/Statistical Methods/Mini-Sentinel PROMPT Group-Sequential-Inverse-Probability-of-Treatment-Weighted-Regression-Tool Users-Guide.pdf).
4. Mini-Sentinel Prospective Routine Observational Monitoring Program Tool: Continuous and Group Sequential Analysis with Self-control Design Technical Users' Guide version 1.0. Mini-Sentinel 2014; available at [http://www.mini-sentinel.org/work\\_products/Statistical Methods/Mini-Sentinel PROMPT Self-Control-Design-Tool Technical-Users-Guide.pdf](http://www.mini-sentinel.org/work_products/Statistical Methods/Mini-Sentinel PROMPT Self-Control-Design-Tool Technical-Users-Guide.pdf).
5. Mini-Sentinel Prospective Routine Observational Monitoring Program Tool: Cohort Matching Technical Users' Guide version 1.0. Mini-Sentinel 2014; available at [http://www.mini-sentinel.org/work\\_products/Statistical Methods/Mini-Sentinel PROMPT Cohort-Matching-Tool Technical-Users-Guide.pdf](http://www.mini-sentinel.org/work_products/Statistical Methods/Mini-Sentinel PROMPT Cohort-Matching-Tool Technical-Users-Guide.pdf).
6. Carnahan R, Gagne JJ, Nelson J, et al. Mini-Sentinel Prospective Routine Observational Monitoring Program Tools (PROMPT): rivaroxaban surveillance. Mini-Sentinel 2014; available at [http://www.mini-sentinel.org/work\\_products/assessments/mini-sentinel\\_prompt\\_rivaroxaban-surveillance-plan.pdf](http://www.mini-sentinel.org/work_products/assessments/mini-sentinel_prompt_rivaroxaban-surveillance-plan.pdf).
7. Prospective routine observational monitoring of mirabegron. Mini-Sentinel 2014; available at [http://www.mini-sentinel.org/work\\_products/Assessments/Mini-Sentinel PROMPT Mirabegron-Surveillance-Plan.pdf](http://www.mini-sentinel.org/work_products/Assessments/Mini-Sentinel PROMPT Mirabegron-Surveillance-Plan.pdf).

8. Gagne JJ, Wang SV, Rassen JA, Schneeweiss S. A modular, prospective, semi-automated drug safety monitoring system for use in a distributed data environment. *Pharmacoepidemiology and drug safety* 2014;23:619-27.
9. Modular Program 3: Frequency of select events during exposure to a drug/procedure group of interest. Mini-Sentinel 2014; available at [http://www.mini-sentinel.org/work\\_products/Data\\_Activities/Mini-Sentinel-Modular\\_Program\\_3-Documentation.pdf](http://www.mini-sentinel.org/work_products/Data_Activities/Mini-Sentinel-Modular_Program_3-Documentation.pdf).
10. Mini-Sentinel Modular Programs Querying Tools: Overview of Functionality and Technical Documentation. Mini-Sentinel 2014; available at [http://www.mini-sentinel.org/work\\_products/Data\\_Activities/Mini-Sentinel-Routine\\_Querying\\_System-Documentation.pdf](http://www.mini-sentinel.org/work_products/Data_Activities/Mini-Sentinel-Routine_Querying_System-Documentation.pdf).
11. Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. *Journal of clinical epidemiology* 2011;64:749-59.
12. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology* 2009;20:512-22.
13. Kulldorff M, Davis RL, Kolczak M, Lewis E, Lieu T, Platt R. A maximized sequential probability ratio test for drug and vaccine safety surveillance. *Sequential Analysis* 2011;30:58-78.
14. Fireman B, Lee J, Lewis N, Bembom O, van der Laan M, Baxter R. Influenza vaccination and mortality: differentiating vaccine effects from bias. *American journal of epidemiology* 2009;170:650-6.