MINI-SENTINEL METHODS

DEVELOPING, IMPLEMENTING, AND TESTING A PROGRAM FOR HIGH-DIMENSIONAL PROPENSITY SCORE ADJUSTMENT IN THE MINI-SENTINEL DISTRIBUTED DATA ENVIRONMENT

Prepared by: Joshua J Gagne, PharmD, ScD,[1] Shirley V Wang, PhD, ScM,[1] Jeremy Rassen, ScD,[1] Nicolas Beaulieu, MA,[2] Jeffrey S Brown, PhD,[2] Sebastian Schneeweiss, MD, ScD[1]

Author Affiliations: 1. Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham & Women's Hospital and Harvard Medical School, Boston, MA 2. Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA

June 2013

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 HHSF223200910006.

Mini-Sentinel Methods

Developing, Implementing, and Testing a Program for High-Dimensional Propensity Score Adjustment in the Mini-Sentinel Distributed Data Environment

Table of Contents

I. EXECUTIVE SUMMARY

Mini-Sentinel is interested in developing the capability to adjust for confounding in rapid assessments. Propensity scores are a particularly useful confounder adjustment technique for between-person comparisons in a distributed data setting as they facilitate simultaneous adjustment for many confounders and preserve data confidentiality. The high-dimensional propensity score algorithm (hd-PS) has been proposed as a method for semi-automated confounder identification and adjustment in routine surveillance activities. However, it was previously unknown whether the hd-PS algorithm faced technical and practical barriers in a distributed environment.

This pilot project investigated the feasibility of incorporating PS and hd-PS adjustment in the Mini-Sentinel distributed environment. The project included the development of a new module that allows users to select a range of PS and hd-PS related parameters and outputs analytic datasets with corresponding tables, figures and other diagnostic information. The new module is fully compatible with the Mini-Sentinel Common Data Model (MSCDM). It has been developed as a standalone module to work in conjunction with Mini-Sentinel Modular Program 3 (Frequency of Select Events During Exposure to a Drug/Procedure Group of Interest), produced by the Mini-Sentinel Operations Center (MSOC), which identifies the cohorts that the PSs and hd-PSs are calculated and used. The propensity score adjustment module was developed using Mini-Sentinel programming guidelines for distributed querying and was designed to be used in a distributed data environment.

The module went through beta-testing with two example drug, event, and predefined confounder scenarios at four Data Partners to identify barriers and solutions for full, rapid implementation. The testing process included selection of a range of parameter settings, creation of modular program packages for distribution, execution of the code at the selected Data Partners, and return of necessary files to the MSOC for analysis and reporting. The module ran successfully at all four Data Partner sites. Feedback from the Data Partners, as well as the logs, diagnostics and related output returned to the MSOC are described in this report. Abbreviated documentation for this module has been developed. The project team worked closely with the 4.10 Active Surveillance Framework Workgroup and the FDA to ensure full compatibility with those activities. Currently, the module is being integrated into the Prospective Routine Observational Monitoring Program Tool (PROMPT): propensity score-matching program of the 4.10 Active Surveillance Framework Workgroup.

II. OBJECTIVE

The objective of this Task Order activity was to develop a scalable modular program to implement predefined propensity score (PS) and high-dimensional propensity score (hd-PS) adjustment within the Mini-Sentinel Distributed Database (MSDD) environment.

III. APPROACH

A. BRIEF METHODS OVERVIEW

1. Propensity scores

Confounding due to differences in baseline outcome risk between treatment groups is a key threat to the validity of Mini-Sentinel assessments.[1] Propensity scores (PSs), which represent patients' probability of receiving a drug of interest conditional on measured baseline variables, are summary scores that help minimize confounding.[2] Variables included in a traditional PS model are selected by the investigator. Within an assessment, patients with the same PS should, on average, have similar distributions of variables used to estimate the score.[2] Balancing treatment groups on the PS therefore balances measured confounders, on average, between groups. PSs are particularly useful in the setting of drug safety monitoring of rare events since they model the exposure (i.e., treatment initiation) rather than outcome and can therefore often include many more potential confounders than traditional outcome modeling.[3] Additionally, by summarizing patients' demographic and clinical characteristics in a single scalar variable, PSs facilitate the multivariable confounding adjustment in a distributed data setting without the need to share identifiable patient information.[4,5]

2. High-dimensional propensity scores (hd-PSs)

The high-dimensional propensity score (hd-PS) algorithm enables semi-automated selection of and adjustment for potential confounders not pre-specified by investigators.[6,7] The algorithm assesses each unique code – ICD-9 diagnosis, CP4-4 procedure code, generic drug name, and so forth – recorded for the patients in the cohort, and determines whether a dichotomous variable indicating the presence or absence of each code may be a confounder. It determines this by assessing the univariate associations between (1) each code in a patient's history and the exposure of interest and (2) each code and the outcome of interest, and uses these associations to calculate the expected amount of confounding bias each code produces. The algorithm then orders the variables according to this estimated expected amount of bias and selects the top *k*-ranked variables, where *k* is a user-specified parameter indicating the number of variables to include in the PS.

B. DEVELOPMENT OF A MODULAR PS AND HD-PS PROGRAM

We incorporated the existing hd-PS SAS macro[8] into a modular program that is compatible with the Mini-Sentinel Common Data Model (MSCDM) and works with version 3.0 of MSOC's Modular Program 3 to:

1. Perform PS adjustment based on user-defined variables such as age, sex, and pre-specified outcome risk factors based on diagnosis codes, procedure codes and prior drug use

2. Perform hd-PS adjustment by automatically identifying and including in the PS a large number of variables that empirically behave like confounders, with options to further adjust for predefined variables and/or health service utilization measures (e.g., number of prior physician visits, hospitalizations, etc.)
3. Generate three PSs using:
    a. Predefined covariates only
    b. Predefined and hd-PS identified covariates
    c. Demographic variables such as age and sex plus hd-PS identified covariates only
4. Automatically apply matching algorithms using the three PSs to the eligible cohort(s) in a pre-specified manner; the matching calipers are 0.01, 0.025, and 0.05 on the PS scale
5. Generate diagnostic tools (output tables and figures) that illustrate the extent to which the adjustment module reduces baseline differences between groups; these additional output tables and figures complement the standard output of Modular Program 3. More specifically, the hd-PS module automatically produces the following tools:
    a. A Data Partner specific "Table 1" that summarizes the characteristics of the analysis cohort(s) based on the pre-defined variables and illustrates covariate balance between groups before and after PS and hd-PS matching
    b. Figures demonstrating overall covariate balance using a summary metric (i.e., Mahalanobis distance) and demonstrating balance on specific covariates using various metrics (e.g., absolute difference and standardized difference)
    c. Figures of the PS and hd-PS distributions including their overlap, which may facilitate decisions about whether patients were comparable enough at baseline to proceed with an analysis; these figures also include the c-statistic, a measure of PS model discrimination

The module was designed to ensure that confidential and proprietary information remains behind the firewalls of each Data Partner, while maintaining flexibility in the analytic approach at the MSOC. The approach to using PSs to condense patient-level data has been deemed to comply with the HIPAA Privacy Rule when applied in the context of a Public Health Authority and when used in a research setting, given that certain criteria are met, as outlined elsewhere.[9] In addition, the module performs multiple PS estimation and matching procedures and collects all necessary diagnostic information in a single query, minimizing the need to submit repeat queries to the Data Partners.

C. IMPLEMENTATION AND TESTING

All code was developed and initially tested using data held by the Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital (BWH) and Harvard Medical School. These data have been transformed into the MSCDM format to ensure compatibility with the Mini-Sentinel Distributed Database.

The BWH team worked closely with the MSOC Infrastructure team to conduct several rounds of internal testing before the program was distributed to select Data Partners for additional beta testing. These rounds of testing were required to ensure that (1) the code and its structure met MS formats and conventions as expected by the Mini-Sentinel programming standards for distributed querying; (2) that the module could be executed by the Data Partner analysts; and (3) that the minimum necessary output without patient identifiers would be returned to MSOC.

The BWH and MSOC teams then implemented and tested the scalability of the modular program in the distributed data setting among four volunteer Data Partners of varying population size (two of small size and two for medium to large size) and operating systems (e.g., Windows vs. Unix):

Data Partners / volunteer programmer analyst:

- Aetna / Yihai Liu
- Kaiser Permanente Northwest (KPNW) / Don Bachman
- Kaiser Permanente Northern California (KPNC) / Jack Hamilton
- Group Health Research Institute (GHRI) / Tyler Ross

The two empirical examples used were:

1. Celecoxib versus non-selective non-steroidal anti-inflammatory drugs (NSAIDs) and risk of myocardial infarction (MI) during calendar year 2009
2. Prasugrel versus clopidogrel and risk of MI during calendar year 2010

The Workgroup members defined all necessary inputs and parameters for each example, including: National Drug Codes (NDCs) to identify each drug of interest, International Classification of Diseases (ICD) codes for the MI outcome, relevant ICD and HCPCS procedure codes for pre-defined covariates, look-back windows for defining new user status and the covariate ascertainment period, the follow-up duration during which to identify MI events, and the calendar time period of analysis. All analyses were implemented as single retrospective assessments.

The example scenarios for beta-testing the modular program were packaged and made available to the volunteer Data Partners using the standard MS approach. That is, all the materials were packaged as .zip files and shared via the Mini-Sentinel secure portal. Each package was labeled with a unique request identifier. The structure of the .zip files for each packaged scenario contained the usual four subfolders of MS data request (i.e., *sasprograms*, *inputfiles*, *dplocal*, and *msoc*) and a fifth one for hd-PS specific materials.

The *sasprograms* folder contained a suite of modular programs; the *inputfiles* folder contained .sas7bdat and Microsoft Excel files with inputs specific to the scenario under investigation; the *hdps* folder held the Pharmacoepidemiology Toolbox developed by the Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School[8] and the toolkit macros produced by the MSOC. When a package was made available to the volunteer Data Partners, the *dplocal* and *msoc* folders were empty. The modular program populated these folders with output at each Data Partner. The *dplocal* folder contained output data with patient identifiers and remained at each respective Data Partner site, behind the Data Partner's firewalls. In contrast, the *msoc* folder became populated with the minimum necessary information needed by MSOC, i.e. all of the output data *without patient identifiers*, metadata on run time, log files and other hd-PS diagnostic output.

The packages sent out for testing were designed to be pre-loaded at the MSOC with relevant input files and module specifications for a requested analysis. Because every package contained the same file structure and was designed to run against the MSCDM, the Data Partner analysts running the package needed to only unzip the file, customize a master SAS program (i.e.,

"00_requestid_call_modular_programs.sas") with local information (e.g., Data Partner identification codes, library pathnames), and run this program. The master call program drew on other modular programs, toolbox macros, and input files contained in the distributed package to produce output data, tables, figures, and diagnostics without further input from Data Partner analysts.

Elements of the Pharmacoepidemiology Toolbox, including the hd-PS algorithm, were implemented as Java programs encapsulated by SAS macros. In order to enhance performance, the volunteer Data Partners were provided with instructions on how to allocate sufficient memory to SAS's java subsystem. The volunteer Data Partners were able to make this change prior to running the distributed beta packages. Workgroup members from the Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School worked to resolve any implementation challenges and answer questions from Data Partner analysts.

IV. RESULTS

Details of the iterative testing process of the beta packages with the volunteer Data Partners, including challenges encountered and their resolutions, are presented in Table 1. Not every Data Partner was able to successfully run each package initially. The Celecoxib package went through three testing iterations with programming changes that included minor bug fixes, removing warnings in the logs, and modifications for small population counts. By the third iteration, the program ran successfully without errors at each Data Partner. The Prasugrel package was implemented recycling the Celecoxib package, and therefore went through only one testing iteration with no changes. After this first iteration, the modular program ran to completion without error in each of the four Data Partners.

Of special interest was the use of Java as programming language for the hd-PS macros. In particular, the hd-PS Java libraries required the allocation of a minimum amount of physical memory. This required the local Data Partner staff to set up the amount of memory available. With special instructions from the BWH and MSOC teams the Data Partners were able to correctly configure their environments. Given that this minimum amount of memory may be quite high, running a package with embedded hd-PS macros could have an impact in shared environments. One Data Partner created a dedicated server in order to isolate the rest of the users from any potential negative effects from running the hd-PS packages.

Tables 2 and 3 provide the run-times for each component of the modular program at each Data Partner Site. Additional outputs are available from the Workgroup.

V. DELIVERABLES

The hd-PS SAS code was successfully developed, tested and imbedded into a modular program. The code is capable of generating three PSs with pre-defined and empirically-identified confounders. This code is available for use within Mini-Sentinel and is being incorporated into the Prospective Routine Observational Monitoring Program Tool (PROMPT): cohort matching program of the 4.10 Active Surveillance Framework Workgroup. An important aspect of the larger PROMPT cohort-matching tool is that it generates de-identified, individual-level data set for each Data Partner in each monitoring period, which permits time-to-event analyses. The data set contains the minimum information required for

central aggregation and analysis by the Mini-Sentinel Operations Center, including a de-identified Data Partner indicator, the monitoring period in which each patient was identified, a variable indicating the patients' index dates, event dates, propensity score values, propensity score matched set numbers, subgroup indicators (when requested), and other subgroup variables (age, sex, and race). The propensity score summarizes the necessary information for confounding adjustment while obscuring detailed patient-level information. This approach has been reviewed by a legal expert who confirmed that it complies with HIPAA.[9] The information requested from each Data Partner meets the minimum necessary standard specified in the Mini-Sentinel Principles and Policies. While the PROMPT cohort-matching tool is currently the only Mini-Sentinel program that uses hd-PS, the hd-PS program itself does not create and store individual-level data sets. The program can be used in other Mini-Sentinel activities that do not create individual-level data sets.

The project activities identified barriers to implementation and lessons learned. One Data Partner could not provide output for the Prasugrel modular program due to the organizational requirement of approving JAVA output release. Additional approval time within certain Data Partners may be needed before hd-PS output can be returned to the MSOC.

VI. RECOMMENDATIONS AND NEXT STEPS

The hd-PS module was run successfully at four volunteer Data Partner sites. During the testing process, the module was fine-tuned to ensure smooth implementation on the different operating systems at each site and code was adjusted based on feedback from Data Partner analysts. The program was able to successfully append analytic datasets produced by the MSOC's Modular Program 3 with an "intention-to-treat" analysis option for outcome identification, predefined PSs and hd-PS, and matched pair identifiers. The program also successfully produced associated tables, figures, and logs at each of the Data Partner sites.

Concerns regarding the memory allocation using Java libraries warrant additional testing to further assess scalability of using hd-PS programs on a routine basis within the MSDD. We are currently working with one Data Partner to allocate additional memory, and are working towards a resolution. This would allow Data Partners to determine whether the extra memory allocation has an impact on other users and whether changes to local environments are required to accommodate a more regular use of the hd-PS programs. In addition, a formal quality check and audit of the Pharmacoepidemiology Toolbox macros may be warranted. We plan to implement the quality control checks as we implement the enhancements in the next year, which themselves will undergo quality control.

The Workgroup has developed a program to facilitate the conduct of rapid, semi-automated, distributed, PS-matched, new user cohort assessments. The hd-PS module is easily scalable as it can be run by multiple Data Partners. The program uses validated methods commonly used in pharmacoepidemiology to address limitations of observational healthcare data and provides Mini-Sentinel with a semi-automated tool to adjust for confounders when conducting routine active surveillance activities.

This adjustment program is now being integrated into a larger modular program (PROMPT: propensity score-matching program) that will enable sequential analyses for routine surveillance as data accrue prospectively within the MSDD. This program will aggregate the analytic cohorts returned from each

Data Partner, for each monitoring period, and produce estimates of unadjusted and adjusted rate differences, unadjusted and adjusted hazard ratios, number needed to treat, attributable fraction, population attributable risk as well as the outputs necessary to run a sequential alerting algorithm (such as maxSPRT[10]). The program will also have an option to use a flexible Health Outcome of Interest macro, which accommodates a more complex algorithm for defining the outcome than Modular Program 3 can currently provide. Each component of the modular program for adjustment is available to Mini-Sentinel and can be integrated into future modular programs.

VII. TABLES AND FIGURES

A. TABLE 1. BETA TESTING PROCESS AND FEEDBACK FROM DATA PARTNERS

Beta Testing Package **(Date and testing scenario)**	**Data Partner**	**Testing Result:**
Jan 22, 2013 Beta 1 - Celecoxib vs. NSAID and risk of MI	DP1	n/a
	DP2	n/a
	DP3	Data Partner had issues with using MP3 v2 beta.
	DP4	n/a
	Action	We modified the latest version of MP3 (3.0) and inserted this version into the package.
Jan 29, 2013 Beta 2 - Celecoxib vs. NSAID and risk of MI	DP1	Start of available data = start of study. Error checking macro aborted the program because there was insufficient data available prior to start of surveillance for ascertaining new user status/covariates. Data Partner Comment: Concerned that the error checking macro printed ERROR: and reason for the error in capital letters to the log.
	DP2	n/a
	DP3	Program ran to completion without errors. Data Partner Comment: Concerned about warnings in logs for adjustment/hdPS macro due to code specific to developers operating system.

Beta Testing Package (Date and testing scenario)	Data Partner	Testing Result:
	Action	Revised code: removed warnings in the log due to code specific to developers operating system; error messages from error checking macro use lower case letters; start and end dates for monitoring are later to allow sufficient lag from start of available data at Data Partners; output figures for propensity score distributions include color, normal distribution and histograms.
Jan 31, 2013 Beta 3 - Celecoxib vs. NSAID and risk of MI	DP1	Program ran to completion without errors.
	DP2	Program stalled at the hdPS analysis because library pathname for .jar file had spaces in it. Re-ran without spaces in pathname. Program ran to completion without errors.
	DP3	n/a
	DP4	Program ran to completion without errors. DP ran on small cut of data. Only 2 new users identified in NSAID group, none in celecoxib. As designed, output datasets and tables have the tag “nomatch” instead of “matched”. Data Partner Comment: Concerned about warning messages in the log for the table creation macro that occurred because of the lack of matches.
	Action	Revised the table creation code to remove warnings if the macro is directed down the “nomatch” path.
Feb 15, 2013 Beta 1 - Prasugrel vs. Clopidogrel and risk of MI	DP1	Program ran to completion without errors.
	DP2	Program ran to completion with 1 error. Warning/error: Occurs at the proc candisc for calculation of M-Distance because it is not possible to calculate a pooled covariance matrix when each exposure group has

Beta Testing Package (Date and testing scenario)	Data Partner	Testing Result:
		only 1 patient. Error does not affect program output.
	DP3	Program ran to completion without errors. Data Partner Comments: Add system option NOQUOTELENMAX to prevent warning in log. The inpatient indicator is not completely accurate at all sites. Some sites convert all ICD9 Procedure codes to CPT4 codes. Other sites may convert some codes. Did not find the 00_ program easy to fill out. It would be much easier if the directory names were pre-filled with relative paths. Did not like that data tables were written/deleted from indata folder.
	DP4	Data Partner Comment: This set of programs needs to go through a security review by internal IT staff. This review was initiated due to the use of JAVA within the SAS code. In particular the JAVA programs need to be (1) approved to be safe to run on local servers and (2) confirmed to not generate any proprietary or confidential information or to not contain any code pointing to external servers to share any kind of data. As a result of this pilot activity no formal output was returned to MSOC by DP4.
	Action	The BWH and MSOC teams implemented all recommended revisions to the code for future production use. A revised version of the hd-PS program will be used by the 4.10 Active Surveillance Framework Workgroup.

B. TABLE 2. TIMING TO RUN MODULE FOR 2 CELECOXIB EXAMPLE IN 4 VOLUNTEER DATA PARTNERS

Timing of celecoxib package							
DP	**Cohort identification**	**Adjustment**				**Table creation**	**Figure creation**
		hd-PS	**Matching**	**Pre-defined variable creation**	**Adjustment total**		
DP1	1 h 26 m 53 s	1 h 20 m 56 s	0 h 00 m 14 s	0 h 59 m 56 s	2 h 21 m 39 s	0 h 01 m 49 s	0 h 02 m 51 s
DP2	1 h 10 m 33 s	1 h 13 m 28 s	0 h 00 m 03 s	1 h 19 m 43 s	2 h 33 m 40 s	0 h 02 m 35 s	0 h 03 m 52 s
DP3	0 h 54 m 08 s	0 h 42 m 37 s	0 h 00 m 12 s	0 h 34 m 15 s	1 h 17 m 29 s	0 h 02 m 53 s	0 h 04 m 28 s
DP4	0 h 00 m 14 s	0 h 00 m 00 s	0 h 00 m 00 s	0 h 00 m 03 s	0 h 00 m 07 s	0 h 02 m 39 s	0 h 03 m 53 s

C. TABLE 3. TIMING TO RUN MODULE FOR 2 PRASUGREL EXAMPLE IN 3 VOLUNTEER DATA PARTNERS

Timing of prasugrel package							
DP	Cohort identification	Adjustment				Table creation	Figure creation
		hd-PS	Matching	Pre-defined variable creation	Adjustment total		
DP1	2 h 42 m 22 s	1 h 13 m 18 s	0 h 00 m 02 s	0 h 57 m 28 s	2 h 11 m 18 s	0 h 01 m 02 s	0 h 01 m 27 s
DP2	1 h 57 m 46 s	1 h 01 m 44 s	0 h 00 m 02 s	1 h 09 m 32 s	2 h 11 m 38 s	0 h 02 m 13 s	0 h 03 m 29 s
DP3	0 h 47 m 32 s	0 h 00 m 02 s	0 h 00 m 02 s	0 h 12 m 13 s	0 h 33 m 09 s	0 h 01 m 05 s	0 h 01 m 28 s

VIII. REFERENCES

1. Gagne JJ, Fireman B, Ryan PB, et al. Design considerations in an active medical product safety monitoring system. Pharmacoepidemiol Drug Saf 2012;21 Suppl 1:32-40.

2. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41-55.

3. 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;158:280-7.

4. Rassen JA, Avorn J, Schneeweiss S. Multivariate-adjusted pharmacoepidemiologic analyses of confidential information pooled from multiple health care utilization databases. Pharmacoepidemiol Drug Saf 2010;19:848-57.

5. 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:S83-9.

6. Rassen JA, Schneeweiss S. Using high-dimensional propensity scores to automate confounding control in a distributed medical product safety surveillance system. Pharmacoepidemiol Drug Saf 2012;21 Suppl 1:41-9.

7. 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.

8. Pharmacoepidemiology Toolbox. 2012. (Accessed at http://www.hdpharmacoepi.org.)

9. Evaluating strategies for data sharing and analyses in distributed data settings. 2012. (Accessed at http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_Methods_Evaluating-Strategies-for-Data-Sharing-and-Analyses.pdf.)

10. Kulldorff M, Davis RL, Kolczak M, Lewis E, Lieu T, Platt R. A maximized sequential probability ratio test for drug and vaccin safety surveillance. Seq Anal 2011;3:58-78.