BIRTH CERTIFICATE DATA MATCHING FOR THE POST-LICENSURE RAPID IMMUNIZATION SAFETY MONITORING (PRISM) PROGRAM:

DEVELOPMENT OF STANDARD FILE STRUCTURES FOR BIRTH AND FETAL DEATH CERTIFICATE DATA AND IMPLEMENTATION OF MATCHING

Prepared by: Susan E. Andrade, ScD,[1] Katie Haffenreffer,[2] Robert Rosofsky,[2] Chunfu Liu, ScD,[3] Hozefa A. Divan, PhD, MSPH,[3] Bolarinwa Ekezue, PhD,[3] Nandini Selvam, PhD, MPH,[3] Alison Tse Kawai, ScD,[2] Linda Pointon, MPhil,[2] Sascha Dublin, MD, PhD[4]

Author Affiliations: 1. Meyers Primary Care Institute (Fallon Community Health Plan, Reliant Medical Group, and University of Massachusetts Medical School), Worcester, MA; 2. Harvard Pilgrim Health Care Institute, Boston, MA; 3. HealthCore, Inc., Wilmington, DE; 4. Group Health Research Institute, Seattle, WA

April 25, 2016

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.

Birth Certificate Data Matching for the Post-Licensure Rapid Immunization Safety Monitoring (PRISM) Program:

Development of Standard File Structures for Birth and Fetal Death Certificate Data and Implementation of Matching

Table of Contents

I. EXECUTIVE SUMMARY

A. OVERVIEW OF PROJECT

The Food and Drug Administration's (FDA) Mini-Sentinel is a pilot program that aims to conduct active surveillance to detect and refine safety signals that emerge for marketed medical products. The Post-Licensure Rapid Immunization Safety Monitoring (PRISM) network is a program for vaccine safety surveillance within the Mini-Sentinel program. It includes national Data Partners, state and city immunization registries, and vaccine safety investigators.

The goals of this Mini-Sentinel task order activity were to: 1) identify states for potential matching of health plan to birth certificate and fetal death report data; 2) gather information about the feasibility and process of conducting matches with birth certificate and fetal death report data in selected states; 3) develop a standard file structure for birth and fetal death data within the Mini-Sentinel Common Data Model (MSCDM); and 4) support one or more PRISM Data Partners to conduct birth certificate data matching and create Mini-Sentinel files with these data. The interim report "Birth Certificate Data Matching for the Post-Licensure Rapid Immunization Safety Monitoring (PRISM) Program: Survey of State and City Departments of Public Health" submitted in November 2012 summarizes the findings of the first two goals of the task order (http://www.mini-sentinel.org/work_products/PRISM/Mini-Sentinel_PRISM_Birth-Certificate-Data-Matching_Survey-State-City-Dept-Public-Health.pdf). This final report summarizes the findings of the final two goals of the task order, describing the development of standard file structures for birth and fetal death data and the process and achievement of linking birth certificate data from state/city departments of public health (DPHs) to one Data Partner's (HealthCore's) data.

B. SUMMARY OF FINDINGS

The workgroup developed a proposed MSCDM birth table including 66 variables providing information on gestational age, prenatal care received, parity, gravidity, maternal smoking status, maternal and paternal age, race, and educational level, congenital malformations and other perinatal outcomes, as well as additional medical and health information documented in the child's birth certificate. The proposed MSCDM fetal death table includes 110 final variables, many similar to those in the proposed birth table, with the addition of 46 variables to capture information on the cause of death.

Overall, 216,623 live born deliveries were identified using the mother's health plan data for all states served by health plans associated with HealthCore. Of the 216,623 live born deliveries, 177,243 (82%) were linked to an infant, ranging from 76% to 91% for the four states selected for birth certificate matching. Overall for all states, 81% of deliveries were linked to an infant using the health plan subscriber number, an additional 1% were linked using name/address matching, and 0.2% were linked using birth certificate data only (not previously found using health plan data alone; ranging from 0 to 1% across the four states selected for birth certificate matching).

Birth certificate data were received and transformed to the proposed MSCDM data file specifications for four states (California, Georgia, Missouri, and Virginia). Of the 87,465 mother-infant pairs identified using the health plan data for these four states, 62,979 (72%) were linked to birth certificate data obtained from the states (range 34% to 97% across states).

C. RECOMMENDATION FOR FUTURE EFFORTS RELATED TO BIRTH CERTIFICATE AND FETAL DEATH REPORT MATCHING

This project demonstrated that incorporating birth certificate data in Mini-Sentinel is feasible for a number of states that include a substantial number of births covered by the respective health plans. However, the proportion of deliveries identified in the health plan data that could be linked to birth certificate data varied widely across states. This variability was likely due to the different algorithms used to link the birth certificate data to the health plan data. Only a small percentage of mother-infant linkages (< 1% of deliveries identified in the health plan data across states selected for birth certificate matching) were identified using birth certificate data only (not previously found using health plan data alone).

Further efforts are ongoing to process data for New York City. Other states (Colorado, Florida, Louisiana, Pennsylvania, Utah) were selected for matching with other PRISM Data Partners as part of the Mini-Sentinel Prospective Surveillance Program, PRISM 2012 Activity 2: Influenza Vaccines and Pregnancy Outcomes. These matching activities will not be included in this report but could be analyzed at a later date. Future efforts might include assessing different deterministic and probabilistic algorithms to optimize the numbers and accuracy of the data linkages. Additional efforts might also include implementation of the linkage of fetal death report data to health plan data.

Given Mini-Sentinel's interest in assessing the safety of medical products in pregnant women, the proposed MSCDM birth and fetal death table file structures should be considered for permanent incorporation into the Mini-Sentinel Common Data Model, to provide information that is important for pregnancy-related activities and that is not captured by administrative health plan data. The efforts to link Data Partner data with vital records are substantial and this effort should be evaluated against the gain for obtaining data elements found in vital records.

II. BACKGROUND

The FDA-sponsored Mini-Sentinel is a pilot program whose aim is to help develop a large-scale active surveillance system to monitor the safety of marketed medical products. The Post-Licensure Rapid Immunization Safety Monitoring (PRISM) system conducts the vaccine-related activities of the Mini-Sentinel. PRISM is a partnership between the FDA, four large national health plans, eight state/city immunization registries, and the Harvard Pilgrim Health Care Institute, which acts as the PRISM Core leadership and management team. PRISM uses computerized administrative data from health plans and computerized vaccine data from state and city immunization registries that are linked and updated on an approximately quarterly schedule.

PRISM investigators are currently building the capacity to study the safety of vaccines in pregnant populations. As part of these efforts, the current workgroup was convened to assess the feasibility of incorporating computerized birth certificate and fetal death report data to provide important information on factors that may affect adverse pregnancy outcomes and birth outcomes. These factors include maternal characteristics, pregnancy-related conditions, and gestational age at birth or fetal death. Incorporating birth certificate and fetal death report data would greatly enhance Mini-Sentinel's capability to study the safety of medical products in pregnant women, as birth and fetal death data contain information not captured by other computerized data sources.

As an initial step in the assessment of the feasibility of birth certificate and fetal death report matching, the Birth Certificate Data Matching for the Post-Licensure Rapid Immunization Safety Monitoring (PRISM) Program workgroup gathered information on the process for obtaining the data, as well as data specifications and availability, from select state and city DPHs. PRISM leaders at the Mini-Sentinel Operations Center (MSOC), FDA and Data Partners identified 20 states to be contacted to obtain information on the feasibility of birth and fetal death data matching. These states were selected because they include relatively high populations of the Data Partners' members; in addition, some of the states already participate in exchanging immunization registry data with the PRISM Data Partners. The workgroup, with input from the PRISM leaders at the MSOC and Data Partners, designed a data collection form to systematically collect information related to the process and requirements for conducting birth certificate and fetal death report data matches in the 20 selected states. The questionnaire and other relevant explanatory documents were e-mailed to the state and city DPHs; the initial contact was followed by additional contacts by e-mail or telephone, as needed. The results of the questionnaire have been previously reported (http://www.mini-sentinel.org/work_products/PRISM/Mini-Sentinel_PRISM_Birth-Certificate-Data-Matching_Survey-State-City-Dept-Public-Health.pdf). The information received from the state/city DPHs from the completed questionnaires was intended to assist the assessment of the feasibility of birth and fetal death data matching by providing: 1) information (data dictionaries) to inform the development of standard file structures for birth and fetal death data, and 2) information about the process for obtaining birth and fetal death data, to assist one Data Partner, HealthCore, to conduct birth certificate data matching with up to 5 states and to create Mini-Sentinel files with these data.

The experience of the FDA-funded Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP) was also used to inform efforts for assessing the feasibility of birth certificate and fetal death report matching. MEPREP is a multisite collaborative research program developed to enable the conduct of studies of medication use and outcomes in pregnancy.[1] Standard data files have been created at each MEPREP site linking healthcare data for women delivering a live-born infant between January 1, 2001 and December 31, 2008 and infants born to these women; these files include maternal and infant characteristics, medication use, and medical care at 11 health plans within 9 states, as well as birth certificate data obtained from the state departments of public health. The birth certificate data obtained from the state DPHs include information on gestational age, prenatal care received, parity, gravidity, maternal smoking status, maternal and paternal age, race, and educational level, congenital malformations and other perinatal outcomes, as well as additional medical and health information documented in the child's birth certificate.

III. METHODS

A. OVERVIEW OF DESIGN

The Mini-Sentinel Birth Certificate Data Matching for PRISM project is a collaboration between the PRISM leaders at FDA and the Mini-Sentinel Operations Center (MSOC), PRISM Data Partners (Aetna, HealthCore, Humana, and Optum) and selected Academic Partners at the Meyers Primary Care Institute, Group Health Research Institute, and Harvard Pilgrim Health Care Institute. This phase of the project entailed: 1) selection of up to 5 states for birth certificate matching with one PRISM Data Partner, HealthCore; 2) development of standard file structures for birth and fetal death data; and 3) linkage of birth certificate data from state/city DPHs to health plan data.

PRISM leaders at the MSOC, FDA, and workgroup members reviewed the MEPREP birth certificate data file specifications and data file formats received by state/city DPHs to determine variables of interest for potential PRISM projects in order to create standard file structures for the MSCDM birth and fetal death tables that would adequately capture information for all states selected for linkage.

With support from other workgroup members, HealthCore investigators and staff completed and submitted applications to request birth certificate data from state/city DPHs. After approval of the applications/requests for access to the data, the HealthCore team worked with the DPH contacts to determine the process for data transfer, file type and content. The workgroup translated and mapped the coding in the data formats received from the state DPHs to the coding in the proposed MSCDM data file specifications. Based upon the mapping specifications, the MSOC developed a set of programs to perform quality checks on the data received from the states and to transform the files for each state into the proposed MSCDM file formats.

B. PROCESS FOR THE SELECTION OF STATES

PRISM leaders at the MSOC and workgroup members aimed to select 5 states for birth certificate data matching with one Data Partner, from the 15 states who returned a completed questionnaire as part of this pilot project. As outlined in the workgroup's interim report (http://www.mini-sentinel.org/work_products/PRISM/Mini-Sentinel_PRISM_Birth-Certificate-Data-Matching_Survey-State-City-Dept-Public-Health.pdf), criteria considered for selection of the 5 states included: 1) the numbers of births within each state for the Data Partner participating in the project; 2) the states for which the Data Partner currently conducts matches with state immunization registry data as part of ongoing PRISM activities; 3) availability of variables of potential interest (e.g., maternal smoking, gestational age, specific congenital anomalies); 4) whether there is preference for the state (vs. Data Partner) to perform the matching; 5) the complexity of the administrative process and expected time to obtain data; 6) availability of recent data; and 7) ability to write birth and fetal death data to a common data format, thus streamlining the work for Data Partners. The cost for obtaining data and/or DPH personnel labor for preparing data files was also taken into consideration. While the numbers of births within the state was a primary consideration, responses to items on the questionnaires completed by the contacts at the state and city departments of public health provided information to address the latter issues.

C. DEVELOPMENT OF STANDARD FILE STRUCTURES FOR BIRTH AND FETAL DEATH DATA

1. Birth Certificate Data

PRISM leaders at the MSOC, FDA, and workgroup members reviewed the Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP) birth certificate data file specifications to determine variables of interest for potential PRISM projects and to use as a template for the standard MSCDM birth certificate file structure. After determination of variables of interest, data file formats received from state/city DPHs considered for selection for conducting linkages (see Section IV.A) were compared to the MEPREP data file specifications (by at least one investigator and the project manager). Potential refinements (i.e., additional variables and/or refinement of coding for existing variables within the MEPREP data file specifications) were noted for discussion with the other workgroup members in order to create a standard file structure for the MSCDM birth table that would adequately capture information for all states selected for linkage.

2. Fetal Death Report Data

The workgroup did not have an existing template for standard file structures for fetal death data file structures (MEPREP does not incorporate fetal death report data). However, for the states of interest that release fetal death data (see Section IV.A), the variables and coding are similar for fetal death and birth data, with additional variables in the fetal death report data that provide information on the cause of fetal death. Similar to the process for the development of the birth certificate file structure (see Section III.C.1 above), PRISM leaders at the MSOC, FDA, and workgroup members determined variables of interest for potential PRISM projects. The workgroup reviewed and compared data file formats received by the states to create a standard file structure for the MSCDM fetal death table.

D. LINKAGE OF BIRTH CERTIFICATE DATA FROM STATE/CITY DPHS TO HEALTH PLAN DATA

1. Application Process

The workgroup obtained applications to request birth certificate data from the DPHs along with completed questionnaires from our initial communications with the states, or from links on the DPH websites, and requested additional information on the process for obtaining data, if necessary. The workgroup also drafted text for common items requested on the applications related to project objectives, description of the project design, analysis plan, and variables being requested. As the actual recipient of the data from states, HealthCore investigators and staff completed and submitted the applications. Workgroup members assisted HealthCore with follow up to the DPH contacts regarding specific DPH questions about the applications and questions for the DPH related to the application process or data, if necessary. A tracking form was developed to document the status of the applications and specific questions and issues that arose during the application process.

2. Identification of Live Born Deliveries and Linkage of Mothers and Infants

PRISM leaders and workgroup members developed a list of International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) diagnosis and procedure codes and Current Procedural Terminology (CPT) codes to identify live born deliveries using health plan data. The list was developed based upon codes used in MEPREP and other prior pregnancy projects and updated to include codes that became available in more recent years (e.g., codes in the ICD-9-CM 640 – 677 range [includes complications related to pregnancy, normal delivery, complications of labor and delivery, complications of puerperium] and V27 [outcome of delivery], and V30-V39 [live born infants according to type of birth]).

The lead PRISM programmer developed a group of SAS programs to identify women ages 10 to 55 years with codes for delivery of a live born infant in the inpatient setting during the period January 1, 2004 to November 30, 2011. Live born delivery codes occurring less than 270 days after the date of a previous live born delivery were not included as new events. In addition, health plan enrollment in a non-Administrative Services Only (ASO) plan was required from 180 days before the pregnancy start date through 30 days after the date of delivery. The HealthCore programmer ran the group of SAS programs on the Mini-Sentinel Distributed Database (MSDD). The information from the MSDD was re-linked to Data Partner's health plan source data (not included in the MSDD) to determine patient names, addresses, facility of delivery, and subscriber numbers, in order to link the mothers to the infants that were delivered, and to create files to transfer to the DPHs, for states where the DPH linked the information to the birth certificate data.

The workgroup developed documentation on the recommended process for matching of mothers to infants identified by the distributed SAS programs (see **Appendix A,** Steps for Mom-Baby Matching Process, Version 7). This process included steps to match deliveries (mothers) and infants using Data Partner subscriber numbers, followed by name-address matching, and birth certificate data, if possible, resulting in creation of a final internal matched file of mothers and infants.

3. Transfer and Linkage of Health Plan Data to Files from the DPH

After approval of the applications/requests for access to the birth certificate data, the HealthCore team worked with the DPH contacts to determine the process for data transfer (e.g., secure file transfer protocol [SFTP]), file type (e.g., SAS, flat text file) and content (e.g., variables necessary for linking the health plan data to the vital records data, for states where the DPH conducted the matching)**.** Information on the transfer process, type and content of data transferred, and method for linking the birth certificate data to the health plan data was documented in a tracking form.

4. Transformation System and Quality Assurance

For the states selected for birth certificate data matching, the workgroup translated and mapped the coding in the state's data formats to the coding in the proposed MSCDM specifications. This included rules for such tasks as simple data transformations/recoding (e.g., convert numeric 1 to character "M" for male gender), parsing (e.g., separate the 10 digits in the single congenital anomalies character string into numeric Boolean values for the 17 individual congenital anomaly variables), and more complex logic (e.g., computing gestational age in weeks, using date of last menses and date of birth).

Based upon the mapping specifications, the MSOC developed, in SAS, a system to perform quality checks on the data received from the states. This system also provided error reports for sharing with states, by individual record, of data values that were inconsistent with values expected based on the DPHs' data dictionaries. The system then transformed the received files for each state into the proposed MSCDM file formats. This system provided efficiencies across the entire project, obviating the need for individual Data Partners to perform programming for each individual state's data. It also ensured uniform adherence to the mapping developed by the workgroup and for populating the MSCDM birth certificate tables.

Additional Data Characterization and Quality Review programs were developed and distributed by the MSOC to be executed by HealthCore for each state. The output was used to assess the degree to which the birth certificate MSCDM tables conform to the table specifications and overall data quality standards (e.g., structurally conforms to the model, no missing months, reasonable trends in data values across and within states).

See **Appendix B** for additional documentation on the State Vital Records Transformation System and Data Quality Requirement Protocol.

E. ANALYSIS

We calculated the proportion of deliveries identified from HealthCore's health plan data that could be linked to an infant using the guidelines specified in the document *Steps for Mom-Baby Matching Process, Version 7*, (**Appendix A**) overall and for individual states selected for linkage to birth certificate data. For states that had provided data by August 2015, we also calculated the proportion of mother-

infant linked pairs that could be matched to a birth certificate, as well as the proportions of unlinked deliveries (mothers) that could be matched to a birth certificate. The birth certificate data for New York City were transmitted to HealthCore in September 2015 and processing of the data is ongoing; thus, these data are not included in this report.

IV. RESULTS

A. SELECTION OF STATES FOR BIRTH CERTIFICATE MATCHING

PRISM leaders at the MSOC, HealthCore, and workgroup members initially selected 4 states (California, Georgia, Indiana, and Virginia) and New York City to conduct birth certificate matching. The selection was primarily based upon evaluation of the states/city with the highest number of births for HealthCore and the states/city for which HealthCore conducts matches with state immunization registry data as part of ongoing projects. The Data Release Committee (DRC) of the Indiana State Department of Health denied the request for data based on lack of resources to complete the request and concerns related to the state's statute regarding disclosure of birth certificate data. New York State was considered as a possible alternate state to conduct matching; however, the New York State Department of Health Vital Records contact indicated that the request for data would not be approved due to New York State public health law restrictions on the release of confidential birth information. Thus, in place of Indiana, Missouri was selected as the final state to conduct birth certificate matching.

B. DEVELOPMENT OF STANDARD FILE STRUCTURES FOR BIRTH AND FETAL DEATH DATA

Appendix C shows the proposed MSCDM birth and fetal death tables. PRISM leaders and the workgroup identified 66 final variables of interest for the MSCDM birth table and 110 final variables of interest for the MSCDM fetal death table. Examples of data elements included in the files are:

- General birth information: infant date of birth, infant sex, gestational age at birth, birth weight, plurality, delivery method
- Maternal and paternal characteristics: race, ethnicity, educational level, age, maternal smoking, maternal alcohol use, marital status, date of last menstrual period, month prenatal care began, number of prenatal care visits, gravidity, parity, height, pre-pregnancy weight, previous preterm infant
- Congenital malformations
- Complications of pregnancy and concurrent illness: diabetes, chronic hypertension, pregnancy-induced hypertension, eclampsia
- Cause of death (fetal death table only)

Data dictionaries (file formats) received from 7 state/city DPHs (California, Georgia, Indiana, Missouri, New York City, New York state, Virginia) were reviewed and compared to the MEPREP data file specifications to help with the development of a standard file structure for the MSCDM; these included states (Indiana, New York state) initially selected for birth certificate matching that were ultimately not matched due to denial of the requests for data by the DPH (see Section IV.A). The comparison to the data dictionaries from the states allowed assessment of the adequacy of the availability and coding of the variables in the MSCDM. Thus, development of the proposed MSCDM birth and fetal death tables was an iterative process that occurred simultaneously with mapping/translation of the coding in the

data formats from the state DPHs to the coding in the proposed MSCDM data file specifications (see Section IV.C.4). Minor adaptation of MEPREP file specifications was necessary to create the MSCDM birth table specifications (e.g., changes in the format of date variables, minor changes in coding and further description of mapping for race variables). The coding of variables in the proposed MSCDM fetal death table is similar to that in the proposed birth table, with the addition of 46 variables to capture information on the cause of death.

The proposed MSCDM birth and fetal death table file structures were submitted to the MSOC for consideration for incorporation into the MSCDM.

C. LINKAGE OF BIRTH CERTIFICATE DATA FROM STATE/CITY DPHS TO HEALTH PLAN DATA

1. Application Process

The HealthCore team submitted applications to DPHs for 4 states (California, Georgia, Indiana, and Virginia) and New York City between March 2013 and June 2013; the request for data was denied by the Indiana State Department of Health (see Section IV.A) in March 2013. A complete application was submitted to Missouri in August 2013. The time from submission of the application to approval of the request (including completion of data use and/or confidentiality agreements) ranged from approximately 1 month to 4 months for the four states (California, Georgia, Missouri, Virginia) that approved applications by the end of September 2013. The New York City Department of Health and Mental Hygiene (DOHMH) denied the application on the grounds that the Data Health Code does not permit disclosure to a non-government entity. The MSOC sent a letter from FDA requesting the NYC DOHMH reconsider this decision given that the MSOC and collaborating institutions are acting under contract with and under the authority of FDA (see **Appendix D**). The application was subsequently resubmitted and approved by the NYC DOHMH.

A common question from many states was related to the need for Institutional Review Board (IRB) oversight of this work, at both the MSOC/PRISM Coordinating Center and state DPHs. To clarify the nature of all Mini-Sentinel activities, the workgroup created a supplemental documentation packet, that included: (1) the Mini-Sentinel Privacy Panel White Paper describing data privacy issues in Mini-Sentinel; (2) a letter from the Department of Health and Human Services' Office for Human Research Protections (OHRP) to the FDA stating that the regulations OHRP administers do not apply to the Sentinel Initiative (OHRP oversees all IRBs); (3) a letter from the FDA to the Mini-Sentinel Principal Investigators stating that Mini-Sentinel is a Sentinel Initiative activity); and (4) a letter from FDA indicating that the PRISM project is a Sentinel/Mini-Sentinel project. This supplemental documentation was included in the application submission packet if the state previously indicated the need for documentation on this matter. In many cases, further discussion was needed, requiring a follow up telephone call with the workgroup and state contacts, to describe in more detail how state data will be used, with respect to the need for IRB oversight at the state DPH. For some states (e.g. California) the application process included two stages, state IRB review of the submission packet and then review by a second office/committee for final approval.

2. Identification of Live Born Deliveries Linkage of Mothers and Infants

Overall, 216,623 live born deliveries were identified using the mother's health plan data for all states served by health plans associated with HealthCore (**Table 1**). Of the 216,623 live born deliveries, 174,343 (81%) were linked to an infant using the subscriber number and an additional 2,407 (1%) were linked

using name/address matching and 0.2% were linked using birth certificate data only (not previously found using health plan data alone; see Section IV.C.3. below). Thus, overall 177,243 (82%) deliveries identified using the mother's health plan data were linked to an infant, ranging from 76% to 91% for the four states for which birth certificate matching has been accomplished (see Section IV.C.3).

3. Transfer and Linkage of Health Plan Data to Files from the DPH

The process for transfer of files differed according to whether: 1) the DPH required HealthCore to send a file of women and infants of interest to the state to conduct matching to the birth certificate data (including mothers linked to an infant and women identified with a delivery code but not linked to an infant); or 2) the DPH preferred to send a file to HealthCore to conduct the matching.

For Georgia, Missouri, and Virginia, the DPH required HealthCore to send a file of women and infants of interest to the state to conduct matching to the birth certificate data. For these states, the HealthCore team created files that included data, with personal identifiers, for: 1) linked mother-infant pairs and 2) mothers/deliveries not linked to an infant. A file including data on children not linked to a mother was also created and sent to the states to assist with efforts to identify additional mother-infant linkages using the birth certificate data (see **Appendix A**). The method of file transfer, type of file, and personal identifiers requested by the states to link the data varied. Transfer methods included encrypted CDs exchanged via postal mail (primary method) and Secure File Transfer Protocol sites. File formats included SAS datasets and flat text files.

The methods employed to link the health plan data to the birth certificate data varied across states. **Table 2** shows the methods for linkage for Georgia, Missouri, and Virginia. For these states, HealthCore received linked files from the DPH within two months after HealthCore transferred the health plan data files to the DPH. After receipt of the files from the states, the HealthCore team processed the files to identify and remove duplicate records and pre-process the received file for the transformation system.

Once the appropriate data security requirements were approved by the state, the California DPH sent files for all births occurring during 2004 through 2011 to HealthCore to conduct the matching. The HealthCore team used some components of the LINKS Record Linkage Package (http://mchp-appserv.cpe.umanitoba.ca/viewConcept.php?conceptID=1029) to determine possible algorithms for matching the HealthCore records with the California records. These algorithms were specified to an internal HealthCore Programming Group that performed the actual linking (
Table 2).

Of the 87,465 deliveries linked to an infant using health plan data in California, Georgia, Missouri, and Virginia (the states for which birth certificate matching has been accomplished), 62,979 (72%) were linked to birth certificate data (ranging from 34% to 97% across states;
Figure 1). Of the 20,860 deliveries not linked to an infant using health plan data, 5,543 (20%) were linked to birth certificate data but did not link to infant health plan data and 493 (2%) were linked to birth certificate data and could also be linked to infant health plan data.

4. Transformation System and Quality Assurance

For the states selected for matching, the number of data dictionaries for birth certificate data that were sent by the states and mapped by the workgroup to the coding in the proposed MSCDM data files specifications (Section IV.B) ranged from 1 for Georgia and Virginia to 8 for California (which used

different file structures for each calendar year of interest). The format and coding of many variables varied considerably across states; these differences were often due to changes to variables with the transition from the 1989 U.S. standard certificate of birth format to adoption of the 2003 U.S. standard certificate format (not all states have adopted the 2003 format).

The HealthCore team used the transformation system described above (Section III.D.4) to perform quality checks on the data received and to transform the files received from the states into the MSCDM format. Minor inconsistencies between the data received and the data dictionaries that were provided by the DPHs were found.

PRISM was quite conservative in managing the inconsistencies in that no assumptions were made as to whether the data were incorrect or whether the data dictionaries were missing information or were incorrect. After following up with the DPH contacts for each state for clarification on the specific variables where inconsistencies were noted, the mapping specifications and look up tables incorporated in the transformation system were revised and used to create the MSCDM birth tables for additional data characterization and quality review. All of the inconsistencies were from identification of data values not found in the data dictionaries.

V. SUMMARY AND RECOMMENDATIONS

This project included the development of standard file structures for birth and fetal death data and the linkage of birth certificate data from state/city departments of public health to health plan data. The workgroup developed a proposed MSCDM birth table including 66 variables providing information on gestational age, maternal and paternal characteristics, congenital malformations and other perinatal outcomes, as well as additional medical and health information documented in the child's birth certificate. The proposed MSCDM fetal death table includes 110 final variables similar to those in the proposed birth table, with the addition of 46 variables to capture information on the cause of death.

Overall, 216,623 live born deliveries were identified using the mother's health plan data for all states served by health plans associated with HealthCore. Of the 216,623 live born deliveries, 174,343 (81%) were linked to an infant using the subscriber number, an additional 2,407 (1%) were linked using name/address matching and 0.2% were linked using birth certificate data only (not previously found using health plan data alone. Birth certificate data were received and transformed to the proposed MSCDM data file specifications for four states (California, Georgia, Missouri, and Virginia). For these states, 72% of the 87,465 deliveries linked to an infant using health plan data were matched to birth certificate data (n=62,979).

The proportion of deliveries linked to an infant in this project (82%) using health plan data was similar to that observed (86%) in 8 health plans collaborating in MEPREP. The MEPREP plans use similar codes to identify deliveries with administrative health data and a similar process to link deliveries to an infant.[2] The proportion of deliveries linked to a birth certificate was lower in this project (72% overall for deliveries linked to an infant in the health plan data) than the proportions of mother-infant pairs linked to birth certificates at the MEPREP sites (95% overall, ranging from 82% to 100% across MEPREP health plans). However, in the present study, the birth certificate linkage rates varied widely across states (range 34% to 97%). This variability was likely due to the different algorithms used to link the birth certificate data to the health plan data. In addition, we used the information on the residence of the

mother to determine the state for the birth certificate data request. Thus, states would not be able to match health plan records to birth certificate data if a woman delivered the infant in a different state than the state in which she resided, or if the state of residence documented in the health plan data was different than the state the woman delivered in for other reasons (e.g. if the woman moved after the date of delivery and the earlier address was not documented in the health plan administrative data). This may partially account for the lower linkage rates in some states, particularly if a high proportion of the members delivered in bordering states.

Further efforts are ongoing to process data for New York City for which matching with HealthCore data was undertaken as part of this feasibility project. An additional consideration for future efforts might include assessing different deterministic and probabilistic algorithms to optimize the numbers and accuracy of the data linkages. Future efforts might also include linkage of fetal death report data to health plan data.

This project demonstrated that incorporating birth certificate data in Mini-Sentinel is feasible for a number of states that include a substantial number of births covered by the respective health plans. Future efforts to incorporate birth certificate data should take into account the challenges that may arise during this process, including the need for sufficient time frames and staff resources to manage the application and data processes. In addition, where feasible, uniform methods should be applied to manage the application content and processes, as well as the data processes. Given Mini-Sentinel's interest in assessing the safety of medical products in pregnant women, the proposed MSCDM birth and fetal death table file structures should be considered for permanent incorporation into the MSCDM, to provide information that is important for pregnancy-related activities and that is not captured by other computerized data sources.

VI. TABLES AND FIGURES

Table 1. Proportion of Deliveries Linked to an Infant

	Total†	California	Georgia	Missouri	Virginia
Deliveries identified in the mother's health plan data	216,623	51,235	22,130	13,351	21,609
Deliveries identified in the mother's data linked to an infant	177,243 (81.8%)	38,835 (75.8%)	19,583 (8.5%)	12,146 (91.0%)	17,394 (80.5%)
Method of linkage*					
Subscriber number	174,343 (80.5%)	38,184 (74.5%)	19,310 (87.3%)	11,972 (89.7%)	16,856 (78.0%)
Name and address matching	2,407 (1.1%)	651 (1.3%)	43 (0.2%)	95 (0.7%)	354 (1.6%)
Linked using birth certificate	493 (0.2%)	0	230 (1.0%)	79 (0.6%)	184 (0.9%)
Deliveries identified in the mother's health plan data not linked to an infant	39,380 (18.2%)	12,400 (24.2%)	2,547 (11.5%)	1,205 (9.0%)	4,215 (19.5%)

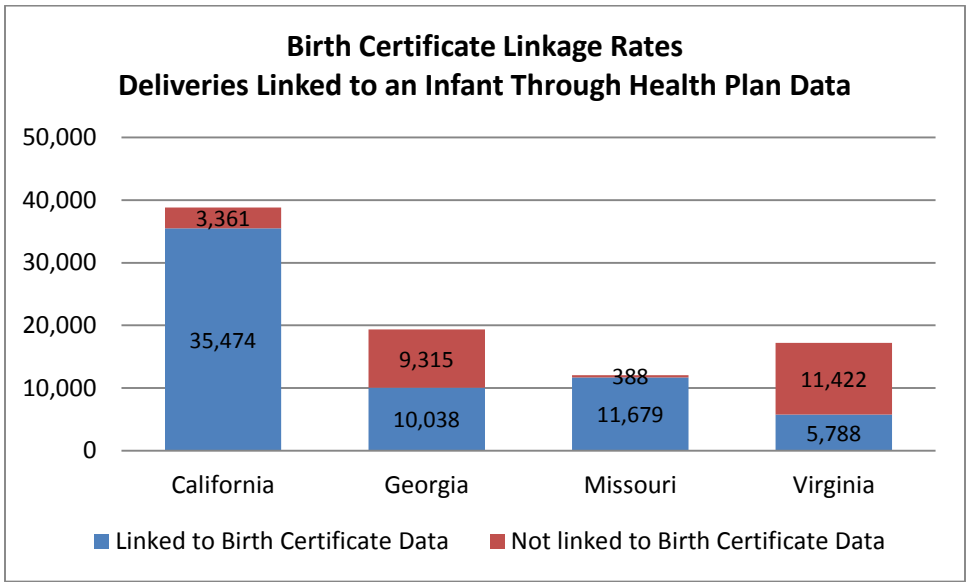
* Methods were applied sequentially, as listed; the denominators for the reported percentages are the total deliveries identified in the mother's health plan data overall and for the respective states.

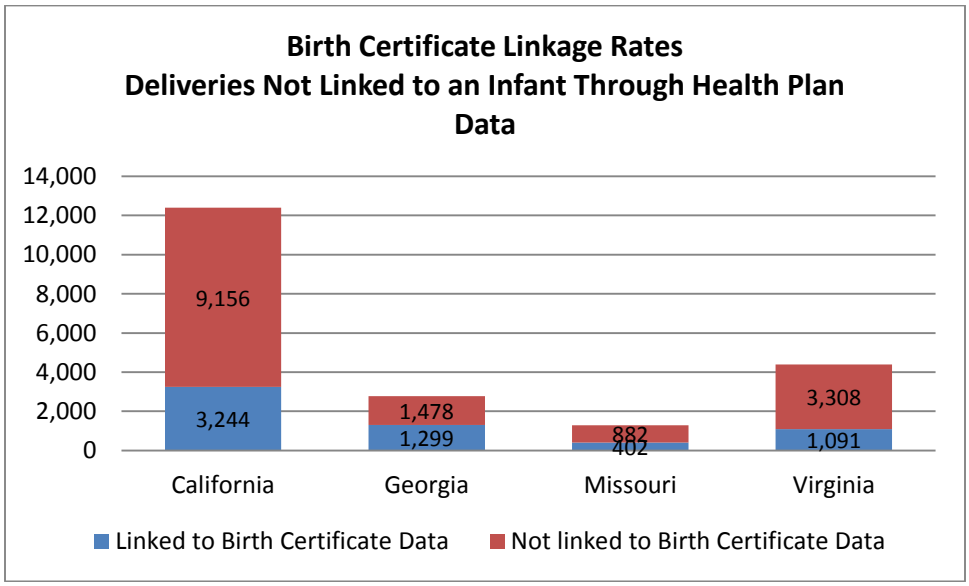
† Total for all states served by health plans associated with HealthCore, including those not selected for potential linkage to birth certificates

Table 2. Methods Used to Link Birth Certificate Data Received From the Departments of Public Health to HealthCore Health Plan Data

	Entity Conducting Matching	Method(s) and Keys Used to Link Health Plan to Birth Certificate Data	Files Received from the State	Patient Identifiers Received from State
California	HealthCore	Deterministic matching 1. child's last name, first name, date of birth, and sex 2. for remaining records from step1, records matched on mother's date of birth, last name, and first name 3. for remaining records from step2, records matched on child's last name, date of birth, and sex	2004 to 2011 births	child's name, child's date of birth, mother's name, mother's date of birth
Georgia	Department of Health	Deterministic matching One of the following: 1. Mother's exact name and date of birth) 2. Soundex of mother's name and date of birth 3. Soundex of mother's name and -/+7 days of date of birth	Separate files for (1) linked mother-infant pairs; 2) mothers/deliveries not linked to an infant	PublicID, event date, facility name, child's name (for linked mother-infant pairs)
Missouri	Department of Health	Deterministic matching 1. Mother's SSN, full name 2. Combination of identifying info such as last name and DOB, first name and DOB, DOB and address, DOB, address, and facility, etc.	Separate linked files for 2004-2009 and 2010-2011	State file number, child's date of birth, child's name
Virginia	Department of Health	Deterministic matching One of the following: 1. Mother's social security number (SSN) 2. Mother's maiden name and age 3. Child's name and date of birth	Separate files of birth certificate data linked by each of the 3 methods	Birth Certificate Number (unique for calendar year), child's name, child's date of birth, mother's SSN, mother's name (those linked by mother's name and age).

Figure 1. Deliveries Linked to Birth Certificate Data Provided by the State Departments of Public Health





VII. ACKNOWLEDGMENTS

The authors would like to thank Grace Lee and Michael Nguyen for their input throughout the project, and Karin Johnson at Group Health Research Institute and Catherine Rogers at the Harvard Pilgrim Health Care Institute for their assistance with mapping the states' data to the coding in the proposed MSCDM specifications.

The workgroup would also like to thank David Cole and Lauren Zichittella for their programming expertise and staff within HealthCore who assisted with the applications to request the data, created the files, and provided the data for the linkages, as well as the contacts at the state and city departments of public health for their assistance in providing information and data for this project.

VIII. REFERENCES

1. Andrade SE, Davis RL, Cheetham TC, et al. Medication Exposure in Pregnancy Risk Evaluation Program. Maternal and child health journal 2012;16:1349-54.
2. Johnson KE, Beaton SJ, Andrade SE, et al. Methods of linking mothers and infants using health plan data for studies of pregnancy outcomes. Pharmacoepidemiology and drug safety 2013;22:776-82.

IX. APPENDICES

A. APPENDIX A. STEPS FOR MOM-BABY MATCHING PROCESS, VERSION 7

PRISM
Mom-Baby Matching Process, Version 7
December 12, 2012

Step 1: Identify Birth Events. Data Partner runs PRISM SAS program to identify all birth events from Mini-Sentinel Common Data Model electronic data. The resulting file is merged with DP source data to create Dataset 1, **Deliveries file**, which includes the following: mother's first and last names, family subscriber number, mother's DOB, date of admission for birth hospitalization, date of discharge, birth facility, and home address.

Step 2: Identify Infants. Data Partner identifies all children with health plan membership born during the years of interest that were enrolled any time (i.e., at least one day) during the first 6 months of life. The resulting file is merged with DP source data to create Dataset 2, the **Infant file**, which includes the following: infant's first and last names, family subscriber number, infant DOB, birth facility (if available) and home address.

Step 3: Subscriber Number Matching. Data Partner links Deliveries file with Infants file based on subscriber number, requiring that the infant DOB falls within the interval of 3 days prior to delivery admission date through delivery discharge date. This generates Dataset 3, the **Subscriber Matches file**, including only those mothers and infants that have been linked in this step. This file should include a flag to indicate that matching was done via subscriber number.

Step 4: Name-Address Matching. For the mothers and infants remaining unlinked after Step 3, Data Partner attempts to link them using the following criteria:

a. baby last name matches mother last name, AND
b. baby address matches mother address,* AND
c. baby's DOB falls within the time window from 3 days prior to mom's admission date for the birth hospitalization through mom's discharge date (3 day window allows for out of hospital births and data anomalies), AND
d. birth facility (if available for both mother and baby. If it is missing for either mom or baby, do not apply this last criterion.)

This generates Dataset 4, **Name-Address Matches**, for only those mothers and infants that have been linked in this step. It should include a flag to indicate that matching was done via name-address.

*Note: We realize that address matching has other complexities, such as the need to standardize addresses (e.g. Road, Rd, etc.) We can provide additional guidance in the future.

Step 5: Data Partner Prepares File for Linking Data Partner Data to State Health Department Data. Data Partner creates a new data file, the **DP Interim Linked File.** This file includes all linked mom-baby

pairs as well as unlinked moms (deliveries) and unlinked infants. Each record should have variables for mom information (such as name, DOB, etc.) and infant information. Some may be left blank for a given record. Specifically:

a. For linked mom-baby pairs: all mother and infant information should be filled in;
b. For unlinked moms (deliveries): only mother information will be included; infant information will be left blank;
c. For unlinked infants: mother information will be left blank; only infant information will be filled in.

This file should not contain any duplicates. That is, if a delivery has been successfully linked to an infant, it should be listed only as a mom-baby pair record and not also included as a delivery-only record.

*At this point, there are **two possible pathways**:*

Pathway 1) The DP sends the DP Interim Linked File to the states to conduct matching to the birth registry, or

Pathway 2) The state sends a file to the DP to conduct matching.

Pathway 1: DP sends file to states to conduct matching to registry. Data Partner sends the DP Interim Linked File to state Departments of Health for linkage to birth certificates.

Step 6: State Departments of Health perform linkage. State Departments of Health perform linkage to the birth registry and return information for successful linkages. Information to be returned to DP may include: the state birth certificate #, birth certificate data for the birth facility, and full names and DOB of mothers and infants. Thus, for example, when a delivery is linked to a birth certificate, the infant's full name and exact DOB are returned. One approach to this might be that the DOH would fill in all blanks on an incomplete record in the **DP Interim Linked File**. So, for example, if the DOH is able to link a delivery-only record to a birth certificate, then the DOH would fill in all the infant information.

Where possible PRISM would like to gather information from States as to how they do matching; for example, gathering information of counts/proportions analogous to Steps 3 and 4.

Pathway 2: States send birth registry file to DP to conduct matching to registry.

Step 7: DOH provides file with identifiers to the DP. Depending on the state, variables may include: mother's SSN, mother's first, maiden, and last name, mother's DOB, birth facility (e.g., facility NPI and/or facility name and address), infant first and last name, infant DOB, infant gender, and birth certificate #.

Step 8: DP matches DP Interim Linked File to DOH file. DP begins by matching linked mom-baby pairs to the state vital records data. Then proceed to link residual unlinked deliveries to the state vital records file and finally the unlinked infants to the file. Details of this process are not provided at this time as the process will vary by state, because of variability in what the state DOH is able to send to the DP.

Where possible, PRISM would like to gather information from DPs as to how they do matching; for example, gathering information of counts/proportions analogous to Steps 3 and 4. Additionally, it would be helpful to identify ordering of matching methods: using mother information, then using baby information, then combinations of information from both, etc.

Common Final Pathway

Step 9: Identifying new linkages using Birth Certificate (BC) Data. This step may not be applicable for some states; it can only be done if it is possible to link delivery-only or infant-only records to the state birth registry. This step could be done by **either the DOH or the DP**. There are multiple possible approaches to this linkage; some examples are listed below.

a. Birth certificate # linkage: Identify previously unlinked deliveries and infants who share the same state birth certificate #. Consider these a linked pair.
b. Name/DOB linkage with delivery as the starting point: For residual deliveries not yet linked to an infant, identify those which linked to a birth certificate. Use the birth certificate information about infant name and infant exact DOB to search for a match within the DP Infant File. (If DOH does this matching, they will be looking within the DP Interim Linked File among the unlinked infants.) Criteria for matching should be similar to step 4 above:
    - infant name in DP file matches that on birth certificate
    - infant DOB in the DP file matches that on the birth certificate
    - If DP conducts this matching, we could also require that infant address in DP file match mother address in DP file. (Note: we have not tried to use address from the birth certificate because families may move, and the address on the birth certificate may be out of date.)
c. Name/DOB linkage with infant as the starting point: Identify infants who have been linked to a birth certificate but not yet to a delivery in the DP file. Use the linked birth certificate to obtain information about the mother including full name, DOB and birth facility. Use this information to search for a matching mother in the DP Deliveries file. (Note: if DOH does the matching, they should seek a match within the unmatched deliveries in the **DP Interim Linked File**.) Matching criteria include:

    - Mother's name in DP file matches mother's name on birth certificate
    - Mother's DOB in DP file matches mother's DOB on infant birth certificate
    - Infant's DOB on the birth certificate file falls within time window of delivery from DP file as described previously
    - If DP conducts this matching, we could also require that infant's address in DP file match mother's address in DP file.

If the DP carries out this work, the final file is the **BC Matches** file, for only those mothers and infants that have been linked in this step. The file should include a flag to indicate that matching was done via birth certificate matching (skip to **Step 11**).

If the DOH does this work, continue to **Step 10**.

Step 10: Ensuring Linkage to DP Delivery and Infant files. If the DOH has done the linking in Step 6, then they will return a file to the DP containing information about deliveries and infants who were not previously known to be linked. The DP will need to link this information back to records within the original Deliveries File and Infant File. The final file of newly linked mom-baby pairs is the **BC Matches** file.

Step 11: Create Final Mom-Baby Pair File. DP creates the **All Matches** file of linked mom-baby pairs. This is the union of the **Subscriber Matches, Name-Address Matches,** and **BC Matches** files. This file should include a flag to indicate which method was used to identify a specific pair. This file should not contain any duplicates (a mom-baby pair should only be recorded once, with the flag for linkage being prioritized as **Subscriber Match > Name-Address Match > BC Match)**.

B. APPENDIX B. STATE VITAL RECORDS TRANSFORMATION SYSTEM AND DATA QUALITY REQUIREMENT PROTOCOL

Mini-Sentinel PRISM
State Vital Records Transformation System
V1.0 – November 25, 2013

Table of Contents

I. Principles

A. All target variables in MSCDM tables are specified in metadata tables.
B. One metadata table specifies variable attributes (i.e., name, type, and length) of MSCDM tables.
C. One metadata table specifies variable names and attributes of State source tables which map to named MSCDM variables. This table also includes information on:
   1. quality control techniques of received data from states
   2. assignment statements of state values to MSCDM values
D. One metadata table specifies variable values for transformations of State source values to MSCDM values. This table is used for strict 1:1 transformation of values.
E. All combinations of states and vital events (i.e., Birth Certificate and Fetal Death) data are included in tables.
F. SAS macro code will generate SAS FORMATs and DATA step assignment statements for identifying data anomalies and SAS code for recoding and transformations.
G. Independent utility macros or user-defined functions will be created for complicated data transformations (e.g., calculating gestational age from date of last menses and date of birth, calculating mother's age from mother's date of birth and child date of birth).
H. System will minimize custom programming required of Data Partners. There likely will be custom programming for Data Partners to accomplish the following:
   1. Preprocess a SAS dataset, from State data returned in a file format of anything other than SAS (e.g., flat file, Access, etc.), ensure uniqueness of records, and include the Data Partner's PatIDs.
   2. Link returned file, from successful matches between Data Partner and State data, with internal Data Partner Mom-Baby matching file or Data Partner Fetal Death Matching file.

II. Processes

A. Quality Control Preprocess for Metadata; developed at MSOC/PRISM, will execute at MSOC/PRISM
   1. Characterize the Birth and Fetal Death MSCDM tables in a metadata table MSCDM_Tables.

2. By State, Vital Event, and Form, ensure that all target variables in MSCDM tables are specified in metadata table Variables_Mapping.
3. By State, Vital Event, and Form, ensure that all target values in MSCDM tables are specified in metadata table Values_Mapping.

B. Quality Control for State and Vital Event files; developed at MSOC/PRISM, will execute at DPs
1. Ensure that all variables named in Variables_Mapping, by State and Vital Event, are in the State/Vital Event dataset. Additional variables in source datasets will be ignored.
2. Ensure that all variables named in Variables_Mapping, by State and Vital Event, have correct source attributes in State/Vital Event dataset. Additional variables in source datasets will be ignored.
3. Ensure that all values in State/Vital Event dataset are found in Values_Mapping. Those not found may require updating of Values_Mapping metadata table or resubmission of file from the State.
4. For variables not named in Values_Mapping, include quality control checking in the Variables_Mapping table.

C. Code Generation; developed at MSOC/PRISM, will execute at DPs
1. Using Values_Mapping, generate SAS formats.
2. Using MSCDM_Tables, generate SAS LENGTH, FORMAT, LABEL, and KEEP statements.
3. Using Values_Mapping and Variables_Mapping, generate SAS assignment statements.
   a) PUT() functions
   b) User-defined functions
   c) Other calculations
4. Create a data set consisting of items I.C.1 through I.C.3

III. Development Steps

A. Create a macro ***Transform***, with the following parameters and preprocessing of parameters as shown:
1. The type of vital event (VitalEvent)
   a) Required
   b) Must be value “B” for birth certificate or “F” for fetal death; no quotes
   c) If not valued, or if not a valid value of “B” or “F”
      Put the following message into the log: “VitalEvent is a required parameter and can be a value of only ‘B’ or ‘F’. Program cannot be processed without a valid value”
      Stop processing the program.
2. The State supplying the source file (State)
   a) Required
   b) Must be any one of the following values: CA, CO, FL, GA, LA, MO, NYC, NC, PA, UT, or VA
   c) If not valued, or if not a valid value shown in b) above.
      Put the following message into the log: “State is a required parameter and can be a value of only CA, CO, FL, GA, LA, MO, NYC, NC, PA, UT, or VA. Program cannot be processed without a valid value”
      Stop processing the program.
3. The version of the form, as a State may have more than one file format for a specific vital event type (Form)
   a) Required

b) Must be an integer valued 1 or higher.
c) If not valued, or if not a valid integer of 1 or higher
   Put the following message into the log: “Form must be valued with an integer value of 1 or higher. Program cannot be processed without a valid value”
   Stop processing the program.

4. The name of the SAS dataset containing the State data (StateFile)
   a) Required
   b) Must be a valid SAS dataset name
   c) If not valued:
      Put the following message into the log: “StateFile is a required parameter. Program cannot be completed without a valid value.”
      Stop processing the program.
   d) If the string is not a valid SAS name:
      Put the following message into the log: “The processed value for the StateFile macro parameter, &StateFile., is not a valid SAS data set name. Program cannot be completed without a valid value.”
      Stop processing the program.
   e) If the parameter is valued and a valid SAS name, but the file does not exist in the proper StateFiles folder:
      Put the following message into the log: “This program cannot be completed unless &StateFile. exists in the StateFiles folder. Please ensure that the parameter is correctly valued and that the file exists.”
      Stop processing the program.
5. The name of the variable that is the State’s unique file number (StateNumberVar)
   a) Required
   b) Must be a valid SAS variable name
   c) If not valued:
      Put the following message into the log: “StateNumberVar is a required parameter. Program cannot be completed without a valid value.”
      Stop processing the program.
   d) If the string is not a valid SAS name:
      Put the following message into the log: “The processed value for the StateNumberVar macro parameter, &StateNumberVar., is not a valid SAS variable name. Program cannot be completed without a valid value.”
      Stop processing the program.
   e) If the parameter is valued and a valid SAS name, but the variable does not exist in the dataset named by the &StateFile parameter:
      Put the following message into the log: “This program cannot be completed unless &StateNumberVar. exists as a variable in the &StateFile. dataset. Please ensure that the parameter is correctly valued and that the variable exists in the file.”
      Stop processing the program.
6. The name of the SAS dataset lookup file containing metadata describing the structure of the MSCDM tables (LookupMSCDM)
   a) Required
   b) If not valued, then will default to “MSCDM_Tables”

c) If the string is not a valid SAS name:
Put the following message into the log: “The processed value for the LookupVariables macro parameter, & LookupVariables., is not a valid SAS data set name. Program cannot be completed without a valid value.”
Stop processing the program.
d) If the resolved file name does not exist in the proper Lookup folder:
Put the following message into the log: “This program cannot be completed unless &LookupVariables. exists in the Lookup folder. Please ensure that the parameter is correctly valued and that the file exists.”
Stop processing the program.

7. The name of the SAS dataset lookup file containing mapping of variables (LookupVariables)
   a) Required
   b) If not valued, then will default to “Variables_Mapping”
   c) If the string is not a valid SAS name:
      Put the following message into the log: “The processed value for the LookupVariables macro parameter, & LookupVariables., is not a valid SAS data set name. Program cannot be completed without a valid value.”
      Stop processing the program.
   d) If the resolved file does not exist in the proper Lookup folder:
      Put the following message into the log: “This program cannot be completed unless &LookupVariables. exists in the Lookup folder. Please ensure that the parameter is correctly valued and that the file exists.”
      Stop processing the program.
8. The name of the SAS dataset lookup file containing mapping of values (LookupValues)
   a) Required
   b) If not valued, then will default to “Values_Mapping”
   c) If the string is not a valid SAS name:
      Put the following message into the log: “The processed value for the LookupValues macro parameter, & LookupValues., is not a valid SAS data set name. Program cannot be completed without a valid value.”
      Stop processing the program.
   d) If the resolved file does not exist in the proper Lookup folder:
      Put the following message into the log: “This program cannot be completed unless &LookupValues. exists in the Lookup folder. Please ensure that the parameter is correctly valued and that the file exists.”
      Stop processing the program.
9. The name of other macros/modules to be called and executed (MCalls)
   a) Required
   b) Must be a space-delimited array of the following allowable abbreviations. They can occur in any order
      - QC_Table_MSCDM: For calling the QC_Table_MSCDM macro
      - QC_Table_Variables: For calling the QC_Table_Variables macro
      - QC_Table_Values For calling the QC_Table_Values macro
      - QC_Table_State: For calling the QC_Table_State macro
      - Generate_Code: For calling the Generate_Code macro

c) If not valued, or not valued with specific abbreviations:
Put the following message into the log: "MCalls must be valued with at least one of only the following values: QC_Table_MSCDM, QC_Table_Variables, QC_Table_Values, QC_Table_State, Generate_Code. Program cannot be processed without valid values."
Stop processing the program.

B. Create a macro ***QC_Table_MSCDM***, called by macro ***Transform*** that will perform QC on the MSCDM_Table. This macro will inherit parameter values from the ***Transform*** macro. It will QC the data per the following specifications:
1. Each MSCDM_Tables row shall have the following:
   - Event is one of "B" or "F"
   - Variable is filled
   - Variable is a valid SAS name
   - Type is one of "N" or "C"
   - Length is integer and greater than zero
   - Format, if filled, is a valid SAS format
   - Label, if filled, is a valid SAS label
2. No duplicate rows in table

C. Create a macro ***QC_Table_Variables***, called by macro ***Transform*** that will perform QC on the &LookupVariables named table. This macro will inherit parameter values from the ***Transform*** macro. It will QC the data per the following specifications:
1. Each Variables table row shall have the following:
   - State is one of identified states (i.e., CA, CO, FL, GA, LA, MO, NYC, NC, PA, UT, or VA)
   - Event is one of "B" or "F"
   - Form is filled with a character digit
   - Year_Start, if filled:
   - Is integer
   - Has a Year_End and vice-versa
   - Year_End, if filled:
   - Is integer
   - If Year_Start and Year_End are filled:
   - Year_Start <= Year_End
   - State_Var is filled.
   - State_Var, if filled, is a valid SAS name
   - MSCDM_Var, if filled, is found in MSCDM_Tables for same value of Event.
   - State_Type is filled with one of "N" or "C"
2. No duplicate rows in table

D. Create a macro ***QC_Table_Values,*** called by macro ***Transform*** that will perform QC on the &LookupValues named table. This macro will inherit parameter values from the ***Transform*** macro. It will QC the data per the following specifications:
1. Each Values table row shall have the following:
   - State is one of identified states (i.e., CA, CO, FL, GA, LA, MO, NYC, NC, PA, UT, or VA)
   - Event is one of "B" or "F"
   - Form is filled with a digit only

- Year_Start, if filled:
  - Is integer
  - Has a Year_End and vice-versa
- Year_End, if filled:
  - Is integer
- If Year_Start and Year_End are filled:
  - Year_Start <= Year_End
- MSCDM_Var is filled and is found in MSCDM_Tables for same value of Event
- State_Var if filled, is found in LookupVariables
- State_Type is filled with one of "N" or "C"
- State_Value is filled with a character value or one of the following:
  - [missing], to indicate allowable missing values
- MSCDM_Value is filled with a character value or one of the following:
  - [missing], to indicate allowable missing values
- If State_Value is filled then MSCDM_Value must be filled and vice-versa
- The pairing of MSCDM_Var and State_var is identical to that in Variables_Mapping for each combination of State, Event, and Form
- Assignment is filled when both State_Value and MSCDM_Value are filled with Blank

2. No duplicate rows in table

E. Create a macro ***QC_Table_State,*** called by macro ***Transform*** that will perform QC on the & StateFile named table. This macro will inherit parameter values from the ***Transform*** macro. It will QC the data per the following specifications:
   1. Ensure that all variables named in Variables_Mapping, by State and Vital Event, are in the State/Vital Event dataset. Additional variables in state source datasets will be ignored.
   2. Ensure that all variables named in Variables_Mapping, by State and Vital Event, have correct source attributes (i.e., numeric or character) in State/Vital Event dataset. Additional variables in source datasets will be ignored.

   For these two purposes, perform the following:
      a) Create a list of variable names and data types from the State file. Assign the values of State, Event, and Form to the list.
      b) Compare the list created in III.E.2.a), to the same list in Variables_Mapping, for each combination of State, Event, and Form.
      c) Write any discrepancies to error file, noting the Variables_Mapping values and the observed values. Then terminate program.
   3. Ensure that all numeric values in State/Vital Event dataset are within the Min and Max values, as named in the Variables_Mapping table where Min and Max values are filled. Those not found may require updating of Values_Mapping metadata table.
      a) Create a list of Min and Max values by variable, from the State file. Assign the values of State, Event, and Form to the list.
      b) Compare the list of values created in III.E.3.a), to the same range of Min and max values in Variables_Mapping, for each combination of State, Event, Form, and State_Var.
      c) Write any discrepancies to error file, noting the Variables_Mapping values and the observed values. Then terminate program.
   4. Custom error checking: Where needed, call user-defined functions to check values in the State file.

a) Call user-defined functions (i.e.., ValQCChar, ValQCNum, ICD10QC) to perform checks for specific sets or patterns of values.
b) Write any discrepancies to error file, noting the Variables_Mapping values and the observed values. Then terminate program.

5. Check that for dataset variables found in Values_Mapping, that all values in the State/Vital Event are found in the lookup table. Perform this check except for the same named variables found in the Variables_Mapping table that have Min and Max values filled. Those values not found in the State/Vital Event file may require updating of Values_Mapping metadata table.
   a) Create a list of values by variable, from the State file. Assign the values of State, Event, and Form to the list.
   b) Compare the list of values created in III.E.5.a), to the same list of values in Values_Mapping, for each combination of State, Event, Form, and State_Var.
   c) Write any discrepancies to error file, noting the Values_Mapping values and the observed values. Then terminate program.
6. For review purposes only, create a report that will consist of the following:
   a) For record identifier variables (e.g., Birth Certificate Number or Fetal Death Number) the value of any identifier that occurs more than once in the file and the number of times it occurs in the file. If no identifier has more than one occurrence then no report will be produced.
   b) For all variables *other* than identifier variables, a report of the variable name, each value for that variable, the count of occurrences of that value in the file, and the proportion of that value occurring in the file. Missing values will be considered a valid value for the purposes of this report.
7. Generate error reports as follows:
   a) For the Data Partner viewing only, create a report, identifying the specific record in which errors occurred and all errors that were found with that record.
   b) For the Data Partner and MSOC viewing, create an aggregate report of discrepancies found, by variable. Include counts and proportions of errors found in the state file.

F. Create a macro ***Generate_Code,*** called by macro ***Transform*** that will generate SAS program code required in order to perform the data transformations on the &StateFile named table. This macro will inherit parameter values from the ***Transform*** macro.
   1. Using Values_Mapping, generate SAS formats. For any combination of State, Event, Form, and MSCDM_Var, where both State_Value and MSCDM_Value are filled as pairs, create a SAS FORMAT
      a) Review the data type for the MSCDM variable. Create a FORMAT for target character data and an INFORMAT for target numeric data.
      b) Name the format using the values of State, Event, Form and MSCDM_Var, separated and suffixed by underscores. For example, for the Birth file from Virginia, where there is only one form, for the ALCOHOL variable, the INFORMAT would be named VA_B_1_ALCOHOL_. Another example for the BSEX variable in the same file, the FORMAT would be named VA_B_1_BSEX_.
      c) Following these processes above, create a CNTLIN dataset, by reading the Values_Mapping table and filtering for specified values of State, Event, and Form.
         - Set fmtname to the value as specified in III.F.1.b)

- Set Type to “C” for a character FORMAT and “I” for a numeric INFORMAT.
- Rename State_Value to Start.
- Rename MSCDM_Value to Label.
- If the State_Value equals “[Missing”], set the Start value to null.
- If the MSCDM_Value equals “[Missing”], set the Label value to null.

d) Always add a row with HLO set to “O” for other.

2. Create a KEEP clause for the MSCDM table, that will filter in only those variables being saved for the table.
   a) Using MSCDM_Tables, for the specified value of Event, create the KEEP clause as follows.
   b) For each row in MSCDM_Tables, for the specified value of Event, output the name of the value specified in MSCDM_Tables.Variable.
3. Create a LENGTH statement for the MSCDM table.
   a) Using MSCDM_Tables, for the specified value of Event, create the LENGTH statement as follows.
   b) Output the “LENGTH” literal.
   c) For each row in MSCDM_Tables, for the specified value of Event, output the name of the value specified in MSCDM_Tables.Variable.
   d) For each row in MSCDM_Tables, if MSCDM_Tables.Type = “C”, output the “$” literal.
   e) For each row in MSCDM_Tables, output the value in MSCDM_Tables.Length.
   f) Output the “;” literal.
   g) NOTE: For variable names that contain the string “PatID” as part of the name, use the maximum length of the PatID variables found among the tables in Data Partner’s analytic library used with this table.
4. Create a FORMAT statement for the MSCDM table.
   a) Using MSCDM_Tables, for the specified value of Event, create the FORMAT statement as follows.
   b) Output the “FORMAT” literal.
   c) For each row in MSCDM_Tables, for the specified value of Event, output the name of the value specified in MSCDM_Tables.Variable.
   d) For each row in MSCDM_Tables, output the value in MSCDM_Tables.Format.
   e) Output the “;” literal.
   f) NOTE: For variable names that contain the string “PatID” as part of the name, use the maximum length of the PatID variables found among the tables in Data Partner’s analytic library used with this table and prefix with “$”.
5. Create a LABEL statement for the MSCDM table.
   a) Using MSCDM_Tables, for the specified value of Event, create the LABEL statement as follows.
   b) Output the “LABEL” literal.
   c) For each row in MSCDM_Tables, for the specified value of Event, output the name of the value specified in MSCDM_Tables.Variable. Follow this with an equals sign and an open double quote.
   d) For each row in MSCDM_Tables, output the value in MSCDM_Tables.Format. Follow this with a close double quote.
   e) Output the “;” literal.

6. Using Variables_Mapping, generate SAS assignment statements. For any combination of State, Event, Form, and MSCDM_Var, create assignment statements as follows:
   a) Variables_Mapping.Assignment is not null, create an assignment statement as:
      Variables_Mapping.MSCDM_Var = Variables_Mapping.Assignment
   b) Variables_Mapping.Assignment is null, process as follows:
      Create assignment statements, using the formats created in step III.F.1 above, following the rules in the following table:

State Data Type	MSCDM Data Type	Format Type	Assignment Statement
Character	Character	Character Format	MSCDM_Var = put(State_Var,character_format_name.)
Numeric	Character	Numeric Format	MSCDM_Var = put(State_Var,numeric_format_name.)
Character	Numeric	Character Informat	MSCDM_Var = input(State_Var,character_format_name.)
Numeric	Numeric	Numeric Informat	MSCDM_Var = input(put(State_Var,#.),numeric_format_name.)

   c) Ensure that any assignment to an MSCDM variable occurs only once, on the basis of either III.F.6.a) or III.F.6.b).
7. Assemble DATA step program.
   a) Write the “DATA” literal, followed by the name of the dataset being created as follows: e_s_f, where “e” is either “Birth” or “FetalDeath” and “s” is the value of State, and “f” is the value of Form. Then follow with a “;” literal.
   b) Include the KEEP clause from III.F.2 above
   c) Include the LENGTH statement from III.F.3.
   d) Include the FORMAT statement from III.F.4.
   e) Include the LABEL statement from III.F.5.
   f) Include each assignment statement from III.F.6.

Data Quality Requirement Protocol

Birth Table

Version 1.0
September 16, 2013

Table of Contents

1 PURPOSE

The purpose of this document is to set the rules associated with data characterization and quality checking of the data contained specifically in the Birth Table of the Mini-Sentinel Distributed Database (MSDD). The MSDD currently contains administrative claims data from participating Mini-Sentinel Data Partners; additional data types will be added in subsequent years. To create the MSDD, each Data Partner transforms their local source data into the Mini-Sentinel Common Data Model (MSCDM) format.

Data checking and characterization is an ongoing process that is undertaken by the MSOC with respect to the entire MSDD, and also a requirement of specific evaluations undertaken by Mini-Sentinel investigators.

The rest of this document defines the Data Characterization and Quality Review process, the specific checks that will be performed and how inconsistencies with the MSCDM will be identified.

1.1 SCOPE

The main driver for the Data Characterization and Quality Review process is to ensure that the MSDD conforms to the MSCDM and that the data included in the MSDD meet reasonable standards for data transformation consistency and quality (e.g., structurally conforms to the model, no missing months, reasonable trends in data values). To evaluate data characteristics and quality, each Data Partner will execute distributed code developed by the MSOC and return aggregated results to the MSOC. MSOC will review results within and across Data Partners and determine if MSCDM and MSDD requirements have been met.

1.2 ASSUMPTIONS AND DEPENDENCIES

The Data Characterization and Quality Review programs must be executed for each update of the Birth table per State. The evaluation will be based on the degree to which the MSDD conforms to the MSCDM and an assessment of overall data quality. The MSOC distributed code executes against the MSDD held by each Data Partner.

2 DATA QUALITY CHECKS

The Data Characterization and Quality Review programs written by MSOC staff and distributed to the Data Partners will be run on the MSDD that should conform to the MSCDM data dictionary format.

2.1 LEVEL 1

Reviews completeness and content of each variable, within the Birth file in the MSDD, to ensure that the required variables are populated and presented in the sequence, content, data type, and formats specified by the MSCDM data dictionary.

2.2 LEVEL 2

Reviews completeness and integrity between any variables within a table, or variables between tables, to identify possible inconsistencies across variables (or tables).

Level 1 and 2 data quality checks generate a Check Summary table as an outcome of running the Data Characterization and Quality Review programs. This Check Summary report is sent to MSOC for review.

2.3 LEVEL 3 DATA QUALITY PROFILING

Level 3 data profiling checks are intended to provide high-level qualitative and quantitative reviews of the data. The data profiling output is a collection of tables containing either frequencies and cross frequencies of many of the data fields of the MSCDM (e.g., frequency of congenital anomalies) or specific data aggregates (e.g., mean value of gestational age).

These tables will be shared with MSOC for quality assurance purposes (e.g., conformity with MSCDM expectations) and also to analyze patterns, trends and characteristics of the data fields. The Level 3 output tables will be used to determine the spread of values within each data field and confirm whether it meets a set of standard expectations. For example, monthly counts of total number of congenital anomalies may help identify an unusual peak associated with duplicated data entries. Similarly, unexpected shifts in the yearly distribution of congenital anomalies by age can reveal inconsistencies in the way the underlying data have been transformed to the MSCDM format over the years. The detail listing of data fields included in the Level 3 are described below.

3 ERROR REPORTING

3.1 ERROR CONDITIONS AND CODES

For all levels of data checks and profiling, acceptable error thresholds are used to identify data errors and anomalies that require discussion with Data Partners. Errors include issues such as nonconformity

with the MSCDM, and anomalies are things such as unusual patterns that require explanation (e.g., large change in enrollment).

Rules and thresholds were determined on a field by field basis. Each rule was associated with a certain threshold and a status. Data quality error statuses include the following:

Status	Status Description
Passed	Data quality check ran without any errors.
Failed	Data quality check ran with fatal errors. Issue must be resolved.
Warning	Data quality check ran without fatal error but minor issues. No action items.

The MSOC reports all identified errors and anomalies to the Data Partners for review and resolution. Resolution can range from documentation of the "anomaly" as valid to recreation of the file (s) as needed.

The tables from Sections **Error! Reference source not found.**, 5 and 6 below list error codes and the rror code descriptions. Acceptable threshold percentages will be developed during Year Two.

3.2 SUMMARY DATA QUALITY REPORT

Using the output tables from the Levels 1-3 data quality check and profiling, a *Summary Data Quality Report* will be generated by MSOC. This will guide Data Partners as to whether any action items remain or if the data meets the set of acceptable thresholds.

4 LEVEL ONE BIRTH TABLE DATA CHECK REQUIREMENTS

The tables list the Level One data checks (by table and by variable).

	Variable Name	Rule	Error Code
1	MPatID	Must be character data type	Bth1.1.1
	MPatID	Must be non-missing	Bth1.1.2
	MPatID	Must be left justified	Bth1.1.3
	MPatID	Must be at least 5 and no more than 100 characters in length	Bth1.1.4
	MPatID	Must include only uppercase letters and/or digits; no embedded blanks or special characters	Bth1.1.5
2	CPatID	Must be character data type	Bth1.2.1
	CPatID	Must be non-missing	Bth1.2.2
	CPatID	Must be left justified	Bth1.2.3
	CPatID	Must be at least 5 and no more than 100 characters in length	Bth1.2.4
	CPatID	Must include only uppercase letters and/or digits; no embedded blanks or special characters	Bth1.2.5
3	State	Must be character data type	Bth1.3.1
	State	Must be non-missing	Bth1.3.2

	Variable Name	Rule	Error Code
	State	Must be left justified	Bth1.3.3
	State	Value must be 2-3 characters in length	Bth1.3.4
	State	Must include only these values: CA CO FL GA LA MO NYC PA UT VA	Bth1.3.5
4	BDOB	Must be numeric data type	Bth1.4.1
	BDOB	Must be of SAS length 4	Bth1.4.2
	BDOB	Must be non-missing	Bth1.4.3
	BDOB	Earliest permitted date is January 1, 2000; latest permitted date is current date	Bth1.4.4
5	BSex	Must be character data type	Bth1.5.1
	BSex	Must be 1 character in length	Bth1.5.2
	BSex	Must be non-missing	Bth1.5.3
	BSex	Must include only uppercase values as follows: “M”, “F”, “A”, or “U”	Bth1.5.4
6	GESTCLIN	Must be numeric data type	Bth1.6.1
	GESTCLIN	Must be 3 bytes in length	Bth1.6.2
	GESTCLIN	Must be positive integers <=45 or missing only	Bth1.6.3
7	GESTMENS	Must be numeric data type	Bth1.7.1
	GESTMENS	Must be 3 bytes in length	Bth1.7.2
	GESTMENS	Must be positive integers <=45 or missing only	Bth1.7.3
8	GESTOBSTET	Must be numeric data type	Bth1.8.1
	GESTOBSTET	Must be 3 bytes in length	Bth1.8.2
	GESTOBSTET	Must be positive integers <=45 or missing only	Bth1.8.3
9	DAYSGEST_OTH	Must be numeric data type	Bth1.9.1
	DAYSGEST_OTH	Must be 3 bytes in length	Bth1.9.2
	DAYSGEST_OTH	Must be missing or integers 112 – 301 inclusive	Bth1.9.3
10	HOW	Must be numeric data type	Bth1.10.1
	HOW	Must be 3 bytes in length	Bth1.10.2
	HOW	Must contain the values 1, 2, 3, -1, or missing only	Bth1.10.3
11	DELIVMETH	Must be character data type	Bth1.11.1
	DELIVMETH	Must be exactly 2 characters in length	Bth1.11.2
	DELIVMETH	Must be left justified	Bth1.11.3
	DELIVMETH	Must include only uppercase values as follows: VA, VB, FO, VC, PC, RC, CU, OT, UK	Bth1.11.4
12	BWEIGHT	Must be numeric data type	Bth1.12.1
	BWEIGHT	Must be 4 bytes in length	Bth1.12.2
	BWEIGHT	Must be positive integers < 8165 or missing only	Bth1.12.3
13	PLURALITY	Must be numeric data type	Bth1.13.1
	PLURALITY	Must be 3 bytes in length	Bth1.13.2
	PLURALITY	Must contain the integer values 1 through 8 or missing only	Bth1.13.3
14	MDOB	Must be numeric data type	Bth1.14.1
	MDOB	Must be SAS length of 4	Bth1.14.2

	Variable Name	Rule	Error Code
	MDOB	Must be non-missing	Bth1.14.3
	MDOB	Earliest permitted date is January 1, 1940	Bth1.14.4
15	MMARSTATUS	Must be character data type	Bth1.15.1
	MMARSTATUS	Must be 1 character in length only	Bth1.15.2
	MMARSTATUS	Must include only uppercase values as follows: M, S, D, W, O, N, OR U	Bth1.15.3
16	MENSDT	Must be SAS length of 4	Bth1.16.1
	MENSDT	Must be non-missing	Bth1.16.2
	MENSDT	Earliest permitted date is January 1, 1940	Bth1.16.3
17	PRENATMON	Must be numeric data type	Bth1.17.1
	PRENATMON	Must be 3 bytes in length	Bth1.17.2
	PRENATMON	Must contain the integer values 0 through 9 or missing only	Bth1.17.3
18	PRENATNUM	Must be numeric data type	Bth1.18.1
	PRENATNUM	Must be 3 bytes in length	Bth1.18.2
	PRENATNUM	Must contain the integer values missing, 0, or positive integers < 200 only	Bth1.18.3
19	MGRAVIDITY	Must be numeric data type	Bth1.19.1
	MGRAVIDITY	Must be 3 bytes in length	Bth1.19.2
	MGRAVIDITY	Must contain the values missing or positive integers only	Bth1.19.3
20	NUM_BTHS	Must be numeric data type	Bth1.20.1
	NUM_BTHS	Must be 3 bytes in length	Bth1.20.2
	NUM_BTHS	Must contain the values missing, zero, or positive integers only	Bth1.20.3
21	PLIV_DEAD	Must be numeric data type	Bth1.21.1
	PLIV_DEAD	Must be 3 bytes in length	Bth1.21.2
	PLIV_DEAD	Must contain the values missing, zero, or positive integers only	Bth1.21.3
22	PLIV_LIV	Must be numeric data type	Bth1.22.1
	PLIV_LIV	Must be 3 bytes in length	Bth1.22.2
	PLIV_LIV	Must contain the values missing, zero, or positive integers only	Bth1.22.3
23	NUM_TRMS	Must be numeric data type	Bth1.23.1
	NUM_TRMS	Must be 3 bytes in length	Bth1.23.2
	NUM_TRMS	Must contain the values missing, zero, or positive integers only	Bth1.23.3
24	MRACE	Must be character data type	Bth1.24.1
	MRACE	Must be 2 characters in length	Bth1.24.2
	MRACE	Must include only values as follows: 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 13, 14, 20, 21, 22, 25, 26, 27, 28, 30, 31, 32, 96, 97, 98, 99, or MU (upper-case only)	Bth1.24.3
25	FRACE	Must be character data type	Bth1.25.1
	FRACE	Must be 2 characters in length	Bth1.25.2

	Variable Name	Rule	Error Code
	FRACE	Must include only values as follows: 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 13, 14, 20, 21, 22, 25, 26, 27, 28, 30, 31, 32, 96, 97, 98, 99, or MU (upper-case only)	Bth1.25.3
26	HISP_MOM	Must be numeric data type	Bth1.26.1
	HISP_MOM	Must be 3 bytes in length	Bth1.26.2
	HISP_MOM	Must contain the values missing or integer values 0 through 5 only	Bth1.26.3
27	HISP_DAD	Must be numeric data type	Bth1.27.1
	HISP_DAD	Must be 3 bytes in length	Bth1.27.2
	HISP_DAD	Must contain the values missing or integer values 0 through 5 only	Bth1.27.3
28	AGE_MOM	Must be numeric data type	Bth1.28.1
	AGE_MOM	Must be 3 bytes in length	Bth1.28.2
	AGE_MOM	Must contain the values missing or positive integers 1-60 only	Bth1.28.3
29	AGE_DAD	Must be numeric data type	Bth1.29.1
	AGE_DAD	Must be 3 bytes in length	Bth1.29.2
	AGE_DAD	Must contain the values missing or integer values 1 through 70 only	Bth1.29.3
30	EDUC_MOM	Must be character data type	Bth1.30.1
	EDUC_MOM	Must be 2 characters in length	Bth1.30.2
	EDUC_MOM	Must include only values as follows: 00, 08, 11, 12, 13, 14, 16, 17, 88, or 99	Bth1.30.3
31	EDUC_DAD	Must be character data type	Bth1.31.1
	EDUC_DAD	Must be 2 characters in length	Bth1.31.2
	EDUC_DAD	Must include only values as follows: 00, 08, 11, 12, 13, 14, 16, 17, 88, or 99	Bth1.31.3
32	TOBACCO	Must be numeric data type	Bth1.32.1
	TOBACCO	Must be 3 bytes in length	Bth1.32.2
	TOBACCO	Must contain the values missing, 0, or 1 only	Bth1.32.3
33	MCIGNUM	Must be numeric data type	Bth1.33.1
	MCIGNUM	Must be 3 bytes in length	Bth1.33.2
	MCIGNUM	Must contain the values missing, 0, positive integers, -88, or -89 only	Bth1.33.3
34	CIGPREPREG	Must be numeric data type	Bth1.34.1
	CIGPREPREG	Must be 3 bytes in length	Bth1.34.2
	CIGPREPREG	Must contain the values missing, 0, positive integers, -88, or -89 only	Bth1.34.3
35	CIGTRIM1	Must be numeric data type	Bth1.35.1
	CIGTRIM1	Must be 3 bytes in length	Bth1.35.2
	CIGTRIM1	Must contain the values missing, 0, positive integers, -88, or -89 only	Bth1.35.3

	Variable Name	Rule	Error Code
36	CIGTRIM2	Must be numeric data type	Bth1.36.1
	CIGTRIM2	Must be 3 bytes in length	Bth1.36.2
	CIGTRIM2	Must contain the values missing, 0, positive integers, -88, or -89 only	Bth1.36.3
37	CIGTRIM3	Must be numeric data type	Bth1.37.1
	CIGTRIM3	Must be 3 bytes in length	Bth1.37.2
	CIGTRIM3	Must contain the values missing, 0, positive integers, -88, or -89 only	Bth1.37.3
38	ALCOHOL	Must be numeric data type	Bth1.38.1
	ALCOHOL	Must be 3 bytes in length	Bth1.38.2
	ALCOHOL	Must contain the values missing, 0, or 1 only	Bth1.38.3
39	DRINKS	Must be numeric data type	Bth1.39.1
	DRINKS	Must be 3 bytes in length	Bth1.39.2
	DRINKS	Must contain the values missing, 0, positive integers, or -88 only	Bth1.39.3
40	WGT_PRE_PREG	Must be numeric data type	Bth1.40.1
	WGT_PRE_PREG	Must be 3 bytes in length	Bth1.40.2
	WGT_PRE_PREG	Must contain the values missing, or positive integers only	Bth1.40.3
41	HGT_MOM	Must be numeric data type	Bth1.41.1
	HGT_MOM	Must be 4 bytes in length	Bth1.41.2
	HGT_MOM	Must contain the values missing or any positive number (can be decimal value) up through 2.13	Bth1.41.3
42	PRV_LT37	Must be numeric data type	Bth1.42.1
	PRV_LT37	Must be 3 bytes in length	Bth1.42.2
	PRV_LT37	Must contain the values missing, 1, or 2 only	Bth1.42.3
43	PRVSMALL_LT37	Must be numeric data type	Bth1.43.1
	PRVSMALL_LT37	Must be 3 bytes in length	Bth1.43.2
	PRVSMALL_LT37	Must contain the values missing, 1, or 2 only	Bth1.43.3
44	NO_CONG_AN	Must be numeric data type	Bth1.44.1
	NO_CONG_AN	Must be 3 bytes in length	Bth1.44.2
	NO_CONG_AN	Must contain the values missing, 1, or 2 only	Bth1.44.3
45	BIFIDA	Must be numeric data type	Bth1.45.1
	BIFIDA	Must be 3 bytes in length	Bth1.45.2
	BIFIDA	Must contain the values missing, 1, or 2 only	Bth1.45.3
46	ANENCEPH	Must be numeric data type	Bth1.46.1
	ANENCEPH	Must be 3 bytes in length	Bth1.46.2
	ANENCEPH	Must contain the values missing, 1, or 2 only	Bth1.46.3
47	MALF_HRT	Must be numeric data type	Bth1.47.1
	MALF_HRT	Must be 3 bytes in length	Bth1.47.2
	MALF_HRT	Must contain the values missing, 1, or 2 only	Bth1.47.3
48	CYAN_CONG_HEART	Must be numeric data type	Bth1.48.1
	CYAN_CONG_HEART	Must be 3 bytes in length	Bth1.48.2

	Variable Name	Rule	Error Code
	CYAN_CONG_HEART	Must contain the values missing, 1, or 2 only	Bth1.48.3
49	OMPHALO	Must be numeric data type	Bth1.49.1
	OMPHALO	Must be 3 bytes in length	Bth1.49.2
	OMPHALO	Must contain the values missing, 1, or 2 only	Bth1.49.3
50	GASTROSCH	Must be numeric data type	Bth1.50.1
	GASTROSCH	Must be 3 bytes in length	Bth1.50.2
	GASTROSCH	Must contain the values missing, 1, or 2 only	Bth1.50.3
51	OMPHALO_UNSP	Must be numeric data type	Bth1.51.1
	OMPHALO_UNSP	Must be 3 bytes in length	Bth1.51.2
	OMPHALO_UNSP	Must contain the values missing, 1, or 2 only	Bth1.51.3
52	TRACH_ESO_FISTUL	Must be numeric data type	Bth1.52.1
	TRACH_ESO_FISTUL	Must be 3 bytes in length	Bth1.52.2
	TRACH_ESO_FISTUL	Must contain the values missing, 1, or 2 only	Bth1.52.3
53	HYPOSPAD	Must be numeric data type	Bth1.53.1
	HYPOSPAD	Must be 3 bytes in length	Bth1.53.2
	HYPOSPAD	Must contain the values missing, 1, or 2 only	Bth1.53.3
54	C_LIP	Must be numeric data type	Bth1.54.1
	C_LIP	Must be 3 bytes in length	Bth1.54.2
	C_LIP	Must contain the values missing, 1, or 2 only	Bth1.54.3
55	PALATE_ONLY	Must be numeric data type	Bth1.55.1
	PALATE_ONLY	Must be 3 bytes in length	Bth1.55.2
	PALATE_ONLY	Must contain the values missing, 1, or 2 only	Bth1.55.3
56	PALATE_UNSP	Must be numeric data type	Bth1.56.1
	PALATE_UNSP	Must be 3 bytes in length	Bth1.56.2
	PALATE_UNSP	Must contain the values missing, 1, or 2 only	Bth1.56.3
57	LIMB_RED	Must be numeric data type	Bth1.57.1
	LIMB_RED	Must be 3 bytes in length	Bth1.57.2
	LIMB_RED	Must contain the values missing, 1, or 2 only	Bth1.57.3
58	DIAPH_HERNIA	Must be numeric data type	Bth1.58.1
	DIAPH_HERNIA	Must be 3 bytes in length	Bth1.58.2
	DIAPH_HERNIA	Must contain the values missing, 1, or 2 only	Bth1.58.3
59	DOWNS	Must be numeric data type	Bth1.59.1
	DOWNS	Must be 3 bytes in length	Bth1.59.2
	DOWNS	Must contain the values missing, 1, 2, or 3 only	Bth1.59.3
60	OTHERCHR	Must be numeric data type	Bth1.60.1
	OTHERCHR	Must be 3 bytes in length	Bth1.60.2
	OTHERCHR	Must contain the values missing, 1, 2, or 3 only	Bth1.60.3
61	DIABETES	Must be numeric data type	Bth1.61.1
	DIABETES	Must be 3 bytes in length	Bth1.61.2
	DIABETES	Must contain the values missing, 1, or 2 only	Bth1.61.3
62	DIAB_GEST	Must be numeric data type	Bth1.62.1
	DIAB_GEST	Must be 3 bytes in length	Bth1.62.2

	Variable Name	Rule	Error Code
	DIAB_GEST	Must contain the values missing, 1, or 2 only	Bth1.62.3
63	DIAB_NONGEST	Must be numeric data type	Bth1.63.1
	DIAB_NONGEST	Must be 3 bytes in length	Bth1.63.2
	DIAB_NONGEST	Must contain the values missing, 1, or 2 only	Bth1.63.3
64	CHYPER	Must be numeric data type	Bth1.64.1
	CHYPER	Must be 3 bytes in length	Bth1.64.2
	CHYPER	Must contain the values missing, 1, or 2 only	Bth1.64.3
65	PIH	Must be numeric data type	Bth1.65.1
	PIH	Must be 3 bytes in length	Bth1.65.2
	PIH	Must contain the values missing, 1, or 2 only	Bth1.65.3
66	ECLAMP	Must be numeric data type	Bth1.66.1
	ECLAMP	Must be 3 bytes in length	Bth1.66.2
	ECLAMP	Must contain the values missing, 1, or 2 only	Bth1.66.3

5 LEVEL TWO: DATA CHECK REQUIREMENTS

Checks are to be performed on non-missing values of variables named, unless otherwise specified.

	Variable Name	Rule	Error Code
1	MPatID	Must have at least one corresponding PatID row in the PRISM Birth Defects VEF “DPLOCAL.PRISMPREG01MOMBABYLINKAGE” table	Bth2.1.1
	MPatID	MPatID variable length must be the same as the PatID length in the PRISM Birth Defects VEF “DPLOCAL.PRISMPREG01MOMBABYLINKAGE” table	Bth2.1.2
2	CPatID	Must have at least one corresponding row in the PRISM Birth Defects VEF “DPLOCAL.PRISMPREG01CHILDLIVEBIRTHMATCH” table	Bth2.2.1
	CPatID	CPatID variable length must be the same as the PatID length in the PRISM Birth Defects VEF “DPLOCAL.PRISMPREG01CHILDLIVEBIRTHMATCH” table	Bth2.2.2
3	BDOB	Fail if absolute values of non-missing BDOB - Demographic.Birth_Date > 30	Bth2.3.1
4	GESTMENS	Fail if absolute value of GESTMENS minus GESTCLIN > 4 weeks	Bth2.4.1
	GESTMENS	Fail if absolute value of GESTMENS minus GESTOBSTET > 4 weeks	Bth2.4.2
5	GESTOBSTET	Fail if absolute value of GESTOBSTET minus GESTCLIN > 4 weeks	Bth2.5.1
6	HOW	If not missing then DAYSGEST_OTH must not be missing	Bth2.6.1

<table>
<tr><th></th><th>Variable Name</th><th>Rule</th><th>Error Code</th></tr>
<tr><td>7</td><td>BWEIGHT</td><td>Compute for PLURALITY=1 and at least one non-missing of GESTCLIN, GESTMENS, and GESTOBSTET.
Take the maximum value of GESTCLIN, GESTMENS, and GESTOBSTET. Flag as failure for any of the following ranges
<table>
<tr><th>Gestational Age</th><th>BWEIGHT >=</th></tr>
<tr><td>< 20</td><td>1000</td></tr>
<tr><td>20-23</td><td>2000</td></tr>
<tr><td>24-27</td><td>3000</td></tr>
<tr><td>28-31</td><td>4000</td></tr>
<tr><td>32-47</td><td>5000</td></tr>
</table></td><td>Bth2.7.1</td></tr>
<tr><td>8</td><td>MDOB</td><td>Fail if MDOB > BDOB</td><td>Bth2.8.1</td></tr>
<tr><td></td><td>MDOB</td><td>Fail if BDOB-Year minus MDOB-Year <=8</td><td>Bth2.8.2</td></tr>
<tr><td></td><td>MDOB</td><td>Fail if BDOB-Year minus MDOB-Year >= 65</td><td>Bth2.8.3</td></tr>
<tr><td></td><td>MDOB</td><td>Fail if absolute values of non-missing MDOB - Demographic.Birth_Date > 30</td><td>Bth2.8.4</td></tr>
<tr><td>9</td><td>MENSDT</td><td>Fail if MENSDT > BDOB</td><td>Bth2.9.1</td></tr>
<tr><td></td><td>MENSDT</td><td>Fail if MENSDT < MDOB</td><td>Bth2.9.2</td></tr>
<tr><td></td><td>MENSDT</td><td>Fail if BDOB minus MENSDT >= 365</td><td>Bth2.9.3</td></tr>
<tr><td></td><td>MENSDT</td><td>Fail if BDOB minus MENSDT < 4 months</td><td>Bth2.9.4</td></tr>
<tr><td>10</td><td>PRENATMON</td><td>Fail if PRENATMON > 0 and either PRENATNUM = 0 or PRENATNUM is missing</td><td>Bth2.10.1</td></tr>
<tr><td>11</td><td>MGRAVIDITY</td><td>If both NUM_BTHS and NUM_TRMS are valued, fail if MGRAVIDITY not equal to sum of NUM_BTHS and NUM_TRMS, plus 1</td><td>Bth2.11.1</td></tr>
<tr><td>12</td><td>NUM_BTHS</td><td>If both PLIV_DEAD and PLIV_LIV are valued, fail if NUM_BTHS not equal to sum of PLIV_DEAD and PLIV_LIV</td><td>Bth2.12.1</td></tr>
<tr><td>13</td><td>AGE_MOM</td><td>Fail if AGE_MOM not equal to calculated age as of BDOB, using MDOB</td><td>Bth2.13.1</td></tr>
<tr><td>14</td><td>TOBACCO</td><td>Fail if TOBACCO=1 and all of MCIGNUM, CIGTRIM1, CIGTRIM2, and CIGTRIM3 equal 0</td><td>Bth2.14.1</td></tr>
<tr><td></td><td>TOBACCO</td><td>Fail if TOBACCO=0 and any non-missing value of MCIGNUM CIGTRIM1, CIGTRIM2, and CIGTRIM3 not equal to 0</td><td>Bth2.14.2</td></tr>
<tr><td>15</td><td>ALCOHOL</td><td>Fail if ALCOHOL=1 and DRINKS = 0</td><td>Bth2.15.1</td></tr>
<tr><td></td><td>ALCOHOL</td><td>Fail if ALCOHOL=2 and non-missing DRINKS not equal to 0</td><td>Bth2.15.2</td></tr>
<tr><td>16</td><td>PRV_LT37</td><td>Fail if PRV_LT37 = 1 and NUM_BTHS = 0 or missing</td><td>Bth2.16.1</td></tr>
<tr><td>17</td><td>PRVSMALL_LT37</td><td>Fail if PRVSMALL_LT37 = 1 and NUM_BTHS = 0 or missing</td><td>Bth2.17.1</td></tr>
<tr><td>18</td><td>DIAB_GEST</td><td>Fail if DIAB_GEST=1 and DIAB_NONGEST=1</td><td>Bth2.18.1</td></tr>
<tr><td>19</td><td>OMPHALO_UNSP</td><td>Fail if OMPHALO_UNSP =1 and either OMPHALO=1 or GASTROSCH=1</td><td>Bth2.19.1</td></tr>
<tr><td>20</td><td>PALATE_ONLY</td><td>Fail if PALATE_ONLY=1 and either PALATE_UNSP=1 or C_LIP=1</td><td>Bth2.20.1</td></tr>
<tr><td></td><td>PALATE_UNSP</td><td>Fail if PALATE_UNSP =1 and either C_LIP=1 or PALATE_ONLY=1</td><td>Bth2.21.1</td></tr>
</table>

6 LEVEL THREE: DATA QUALITY PROFILING

Level 3 checks enable comparisons of descriptive statistics (e.g., trends, distributions) for each variable. The primary comparisons are within and across Data Partner for an individual data extract. Level 3 checks are described below.

This section describes the aggregate file structures to be created from the Data Partner Birth table. This report will be run after each creation and update of the table from a State. The report will provide data-level summary information for each file being submitted per Data Partner (or per state).

For any table that displays statistics by year, the year of birth (from variable BDOB) will be used.

1. Total Birth Records

   For each Data Partner and Totals

	California	Colorado	Florida	*etc.*	Total
Number of Birth Records	176,264	924,602	291,079		3,599,459
Number matched to both a DP mother and baby	109,080	505,989	189,334		2,876,223
Number matched to only a DP mother	42,223	121,207	45,354		176,902
Number matched to only a DP baby	24,961	297,406	56,391		546,334

Note:
Number matched to both a DP mother and baby = Both MPatID and CPatID are filled
Number matched to only a DP mother = Only MPatID is filled
Number matched to only a DP baby = Only CPatID is filled

NOTE: For all subsequent tables, create a new dimension consisting of these 4 levels:

- Birth record matched to both a DP mother and baby (Both MPatID and CPatID are filled)
- Birth record matched to only a DP mother (Only MPatID is filled)
- Birth record matched to only a DP baby (Only CPatID is filled)
- All (All rows)

Then generate separate tables for each of the 4 levels of the dimension.

2. Birth Records By Year
   For each Data Partner and Totals
   (Some blank cells expected as not all DPs will match with all states)

	California	Colorado	Florida	*etc.*	Total
2006	99,999	99,999	99,999	99,999	99,999
2007	99,999	99,999	99,999	99,999	99,999
2008	99,999	99,999	99,999	99,999	99,999
2009	99,999	99,999	99,999	99,999	99,999
2010	99,999	99,999	99,999	99,999	99,999
2011	99,999	99,999	99,999	99,999	99,999

3. Birth Record Report By Year and Gender
   For each Data Partner and Totals
   (Some blank cells expected. Percents are column percents within State/Total)

		California		Colorado		Florida		*etc.*		Total	
Birth Year		**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
2007	Female	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Male	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Ambiguous	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Unknown	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Total	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0
2008	Female	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Male	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Ambiguous	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Unknown	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Total	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0
Total	Female	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Male	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Ambiguous	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Unknown	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Total	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0

NOTE: Use variable BSEX

4. Gestational Age by Year
   For each Data Partner and Totals
   (Some blank rows expected, due to state not providing data)

		California						*etc.*	Total					
Birth Year		**N**	**Mean**	**STD**	**Med**	**Missing N (%)**	**Total N**		**N**	**Mean**	**STD**	**Med**	**Missing N (%)**	**Total N**
2007	GESTCLIN (wks)	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999
	GESTMENS (wks)	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999
	GESTOBSTET (wks)	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999
	DAYSGEST_OTH	99	999.9	99.9	999	99 (99.9)	999		99	999.9	99.9	999	99 (99.9)	999
2008	GESTCLIN (wks)	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999

		California						*etc.*	Total					
Birth Year		**N**	**Mean**	**STD**	**Med**	**Missing N (%)**	**Total N**		**N**	**Mean**	**STD**	**Med**	**Missing N (%)**	**Total N**
	GESTMENS (wks)	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999
	GESTOBSTET (wks)	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999
	DAYSGEST_OTH	99	999.9	99.9	999	99 (99.9)	999		99	999.9	99.9	999	99 (99.9)	999
Etc.	GESTCLIN (wks)	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999
	GESTMENS (wks)	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999
	GESTOBSTET (wks)	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999
	DAYSGEST_OTH	99	999.9	99.9	999	99 (99.9)	999		99	999.9	99.9	999	99 (99.9)	999
Total	GESTCLIN (wks)	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999
	GESTMENS (wks)	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999
	GESTOBSTET (wks)	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999
	DAYSGEST_OTH	99	999.9	99.9	999	99 (99.9)	999		99	999.9	99.9	999	99 (99.9)	999

NOTE: Use variable names specified in table

5. Delivery Method by Year
   For each Data Partner and Totals
   (Some blank cells expected. Percents are column percents within State/Year)

		California		Colorado		Florida		*etc.*		Total	
Birth Year		**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
2007	Repeat C-Section	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Primary C-Section	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	C-Section Unspecified	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Vacuum	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Forceps	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	VBAC	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Vaginal	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Other	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Unknown	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0
2008	Repeat C-Section	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Primary C-Section	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	C-Section Unspecified	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Vacuum	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Forceps	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	VBAC	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Vaginal	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Other	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Unknown	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0
Etc.	Repeat C-Section	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9

		California		Colorado		Florida		*etc.*		Total	
Birth Year		**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
	Primary C-Section	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	C-Section Unspecified	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Vacuum	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Forceps	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	VBAC	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Vaginal	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Other	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Unknown	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0
Total	Repeat C-Section	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Primary C-Section	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	C-Section Unspecified	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Vacuum	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Forceps	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	VBAC	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Vaginal	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Other	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Unknown	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0

NOTE: Use variable DELIVMETH

6. Birthweight by Year
   For each Data Partner and Totals

	California						Colorado						*etc.*	Total					
Birth Year	**N**	**Mean**	**STD**	**Med**	**Missing N (%)**	**Total N**	**N**	**Mean**	**STD**	**Med**	**Missing N (%)**	**Total N**		**N**	**Mean**	**STD**	**Med**	**Missing N (%)**	**Total N**
2007	99	9999.9	99.9	9999	99 (99.9)	999	99	9999.9	99.9	9999	99 (99.9)	999		99	9999.9	99.9	9999	99 (99.9)	999
2008	99	9999.9	99.9	9999	99 (99.9)	999	99	9999.9	99.9	9999	99 (99.9)	999		99	9999.9	99.9	9999	99 (99.9)	999
Etc.	99	9999.9	99.9	9999	99 (99.9)	999	99	9999.9	99.9	9999	99 (99.9)	999		99	9999.9	99.9	9999	99 (99.9)	999
Total	99	9999.9	99.9	9999	99 (99.9)	999	99	9999.9	99.9	9999	99 (99.9)	999		99	9999.9	99.9	9999	99 (99.9)	999

NOTE: Use variable BWEIGHT

7. Plurality
   For each Data Partner and Totals

California						Colorado						*etc.*	Total					
N	Mean	STD	Med	Missing N (%)	Total N	N	Mean	STD	Med	Missing N (%)	Total N		N	Mean	STD	Med	Missing N (%)	Total N
99	9.9	99.9	9	99 (99.9)	999	99	9.9	99.9	9	99 (99.9)	999		99	9.9	99.9	9	99 (99.9)	999

NOTE: Use variable PLURALITY

8. Plurality by Year
   For each Data Partner and Totals

		California		Colorado		*Etc.*		Total	
Birth Year		N	%	N	%	N	%	N	%
2007	Singleton	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Multiple	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Total	99,999	100.0	99,999	100.0	99,999	100.0	99,999	100.0
2008	Singleton	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Multiple	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Total	99,999	100.0	99,999	100.0	99,999	100.0	99,999	100.0
Etc.	Singleton	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Multiple	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Total	99,999	100.0	99,999	100.0	99,999	100.0	99,999	100.0
Tota	Singleton	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Multiple	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Total	99,999	100.0	99,999	100.0	99,999	100.0	99,999	100.0

NOTE: Use variable PLURALITY:
Singleton = (PLURALITY=1)
Multiple = (PLURALITY $\geq$ 2)

9. Marital Status
   For each Data Partner and Totals
   (Percents are column percents within State/Total)

	California		Colorado		Florida		*etc.*		Total	
	N	%	N	%	N	%	N	%	N	%
Married	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
Single	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
Divorced	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
Widowed	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9

	California		Colorado		Florida		*etc.*		Total	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Other	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
Not Married	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
Unknown	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0

NOTE: Use variable MMARSTATUS

10. Month Prenatal Care Began

For each Data Partner and Totals

	California						Colorado						*etc.*	Total					
	N	Mean	STD	Med	Missing N (%)	Total N	N	Mean	STD	Med	Missing N (%)	Total N		N	Mean	STD	Med	Missing N (%)	Total N
Month Prenatal Care Began	99	99.9	99.9	99	99 (99.9)	999	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999
No Prenatal Care						99						99							99

NOTE: Use variable PRENATMON

11. Number of prenatal care visits

For each Data Partner and Totals

	California								Colorado								*etc.*	Total							
	N	Min	Max	Mean	STD	Med	Missing N (%)	Total N	N	Min	Max	Mean	STD	Med	Missing N (%)	Total N		N	Min	Max	Mean	STD	Med	Missing N (%)	Total N
Number of prenatal care visits	99	99	99	99.9	99.9	99	99 (99.9)	999	99	99	99	99.9	99.9	99	99 (99.9)	999		99	99	99	99.9	99.9	99	99 (99.9)	999
No Prenatal Care								99								99									99

NOTE: Use variable PRENATNUM

12. Gravidity (total pregnancies)
    For each Data Partner and Totals

California						Colorado						*etc.*	Total					
N	Mean	STD	Med	Missing N (%)	Total N	N	Mean	STD	Med	Missing N (%)	Total N		N	Mean	STD	Med	Missing N (%)	Total N
99	99.9	99.9	99	99 (99.9)	999	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999

NOTE: Use variable MGRAVIDITY

13. Previous live births total
    For each Data Partner and Totals

California						Colorado						*etc.*	Total					
N	Mean	STD	Med	Missing N (%)	Total N	N	Mean	STD	Med	Missing N (%)	Total N		N	Mean	STD	Med	Missing N (%)	Total N
99	99.9	99.9	99	99 (99.9)	999	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999

NOTE: Use variable NUM_BTHS

14. Parity by Year (Nulliparous vs. Parous)
    For each Data Partner and Totals
    (Percents are column percents within State/Year)

		California		Colorado		*etc.*		Total	
Birth Year	Parity	N	%	N	%	N	%	N	%
2007	Nulliparous	999	99.9	999	99.9	999	99.9	999	99.9
	Parous	999	99.9	999	99.9	999	99.9	999	99.9
	Missing	999	99.9	999	99.9	999	99.9	999	99.9
	Total	999	100.0	999	100.0	999	100.0	999	100.0
2008	Nulliparous	999	99.9	999	99.9	999	99.9	999	99.9
	Parous	999	99.9	999	99.9	999	99.9	999	99.9
	Missing	999	99.9	999	99.9	999	99.9	999	99.9
	Total	999	100.0	999	100.0	999	100.0	999	100.0
Etc.	Nulliparous	999	99.9	999	99.9	999	99.9	999	99.9
	Parous	999	99.9	999	99.9	999	99.9	999	99.9
	Missing	999	99.9	999	99.9	999	99.9	999	99.9
	Total	999	100.0	999	100.0	999	100.0	999	100.0
Total	Nulliparous	999	99.9	999	99.9	999	99.9	999	99.9
	Parous	999	99.9	999	99.9	999	99.9	999	99.9
	Missing	999	99.9	999	99.9	999	99.9	999	99.9
	Total	999	100.0	999	100.0	999	100.0	999	100.0

NOTE: Use variable NUM_BTHS:
0 = Nulliparous
1+ = Parous

15. Previous live births now dead
    For each Data Partner and Totals

California						Colorado						*etc.*	Total					
N	Mean	STD	Med	Missing N (%)	Total N	N	Mean	STD	Med	Missing N (%)	Total N		N	Mean	STD	Med	Missing N (%)	Total N
99	99.9	99.9	99	99 (99.9)	999	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999

NOTE: Use variable PLIV_DEAD

16. Previous live births now living
    For each Data Partner and Totals

California						Colorado						*etc.*	Total					
N	Mean	STD	Med	Missing N (%)	Total N	N	Mean	STD	Med	Missing N (%)	Total N		N	Mean	STD	Med	Missing N (%)	Total N
99	99.9	99.9	99	99 (99.9)	999	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999

NOTE: Use variable PLIV_LIV

17. Previous other pregnancy outcomes
    For each Data Partner and Totals

California						Colorado						*etc.*	Total					
N	Mean	STD	Med	Missing N (%)	Total N	N	Mean	STD	Med	Missing N (%)	Total N		N	Mean	STD	Med	Missing N (%)	Total N
99	99.9	99.9	99	99 (99.9)	999	99	99.9	99.9	99	99 (99.9)	999		99	99.9	99.9	99	99 (99.9)	999

NOTE: Use variable NUM_TRMS

18. Parent Race
    For each Data Partner and Totals
    (Percents are column percents within State/Year)

		California				...				Total			
		Mother		Father		Mother		Father		Mother		Father	
Birth Year	Race	N	%	N	%	N	%	N	%	N	%	N	%
2007	White	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Black	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	...	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Unknown	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Multi-racial	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Totals	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0
2008	White	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Black	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	...	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Unknown	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9

		California				...				Total			
	Multi-racial	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Totals	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0
Etc.	White	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Black	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	...	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Unknown	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Multi-racial	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Totals	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0
Total	White	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Black	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	...	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Unknown	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Multi-racial	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
	Totals	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0

NOTE: Use variables MRACE and FRACE

19. Parent Hispanic Indicator

For each Data Partner and Totals

	California				...				Total			
	Mother		Father		Mother		Father		Mother		Father	
	N	%	N	%	N	%	N	%	N	%	N	%
Non-Hispanic or No	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
Mexican	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
Puerto Rican	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
Cuban	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
...	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
Missing	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
Totals	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0

NOTE: Use variables HISP_MOM and HISP_DAD

20. Parent Age

For each Data Partner and Totals

	California								Colorado								*etc.*	Total							
	N	**Min**	**Max**	**Mean**	**STD**	**Med**	**Missing N (%)**	**Total N**	**N**	**Min**	**Max**	**Mean**	**STD**	**Med**	**Missing N (%)**	**Total N**		**N**	**Min**	**Max**	**Mean**	**STD**	**Med**	**Missing N (%)**	**Total N**
Mother Age	99	99	99	99.9	99.9	99	99 (99.9)	999	99	99	99	99.9	99.9	99	99 (99.9)	999		99	99	99	99.9	99.9	99	99 (99.9)	999
Mother Age Missing/Unknown N (%)								99 (%)								99 (%)									99 (%)
Father Age	99	99	99	99.9	99.9	99	99 (99.9)	999	99	99	99	99.9	99.9	99	99 (99.9)	999		99	99	99	99.9	99.9	99	99 (99.9)	999
Father Age Missing/Unknown N (%)								99 (%)								99 (%)									99 (%)

NOTE: Use variables AGE_MOM and AGE_DAD

21. Parent Education

For each Data Partner and Totals

(Percents are column percents within State/Parent)

	California				...				Total			
	Mother		Father		Mother		Father		Mother		Father	
	N	%	N	%	N	%	N	%	N	%	N	%
No Education	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
8th grade or less	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
9th grade through 12th grade, no diploma	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
High school graduate or GED completed	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
Some college credit but no degree	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
Associate's Degree	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
Bachelor's Degree	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
Education beyond Bachelor's Degree	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
One or more years of college, not stated whether degree earned	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
Missing; unknown	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9	9,999	99.9
Totals	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0	9,999	100.0

NOTE: Use variables EDUC_MOM and EDUC_DAD

22. Tobacco Use

For each Data Partner and Totals

(Percents are column percents within State/Year)

		California		Colorado		Florida		*etc.*		Total	
Birth Year		**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
2007	Yes	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	No	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0
2008	Yes	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	No	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0
Etc.	Yes	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	No	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0
Total	Yes	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	No	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0

NOTE: Use variable TOBACCO

23. Cigarettes

For each Data Partner and Totals

		California			Colorado			*etc.*			Total		
		Mean	STD	Median	Mean	STD	Median	Mean	STD	Median	Mean	STD	Median
During pregnancy	None N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Less than one, occasional use N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Chewing Tobacco N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Missing N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Count	99.9	9.9	99	99.9	9.9	99	99.9	9.9	99	99.9	9.9	99
Pre-pregnancy	None N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Less than one, occasional use N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Chewing Tobacco N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Missing N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Count	99.9	9.9	99	99.9	9.9	99	99.9	9.9	99	99.9	9.9	99
1st Trimester	None N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Less than one, occasional use N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Chewing Tobacco N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Missing N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Count	99.9	9.9	99	99.9	9.9	99	99.9	9.9	99	99.9	9.9	99
2nd Trimester	None N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Less than one, occasional use N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Chewing Tobacco N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Missing N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Count	99.9	9.9	99	99.9	9.9	99	99.9	9.9	99	99.9	9.9	99
3rd Trimester	None N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Less than one, occasional use N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Chewing Tobacco N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Missing N (%)			99 (%)			99 (%)			99 (%)			99 (%)
	Count	99.9	9.9	99	99.9	9.9	99	99.9	9.9	99	99.9	9.9	99

NOTE: Use variables MCIGNUM, CIGPREPREG, CIGTRIM1, CIGTRIM2, and CIGTRIM3

24. Alcohol Use

For each Data Partner and Totals

(Percents are column percents within State)

	California		Colorado		Florida		*etc.*		Total	
	N	%	N	%	N	%	N	%	N	%
Yes	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
No	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0

NOTE: Use variable ALCOHOL

25. Drinks

For each Data Partner and Totals

	California			Colorado			*etc.*			Total		
	Mean	STD	Median	Mean	STD	Median	Mean	STD	Median	Mean	STD	Median
None N (%)			99 (%)			99 (%)			99 (%)			99 (%)
Less than one, occasional use N (%)			99 (%)			99 (%)			99 (%)			99 (%)
Missing N (%)			99 (%)			99 (%)			99 (%)			99 (%)
Count	99.9	9.9	99	99.9	9.9	99	99.9	9.9	99	99.9	9.9	99

NOTE: Use variable DRINKS

26. Maternal weight pre-pregnancy

For each Data Partner and Totals

	California								Colorado								*etc.*	Total							
	N	Min	Max	Mean	STD	Med	Missing N (%)	Total N	N	Min	Max	Mean	STD	Med	Missing N (%)	Total N		N	Min	Max	Mean	STD	Med	Missing N (%)	Total N
Maternal weight	99	999	999	999.9	99.9	999	99 (99.9)	999	99	999	999	999.9	99.9	999	99 (99.9)	999		99	999	999	999.9	99.9	999	99 (99.9)	999
Missing/Unknown N (%)							99 (%)								99 (%)									99 (%)	

NOTE: Use variable WGT_PRE_PREG

27. Maternal height in meters

For each Data Partner and Totals

	California								Colorado								*etc.*	Total							
	N	Min	Max	Mean	STD	Med	Missing N (%)	Total N	N	Min	Max	Mean	STD	Med	Missing N (%)	Total N		N	Min	Max	Mean	STD	Med	Missing N (%)	Total N
Maternal height	99	9.9	9.9	9.9	9.9	9.9	99 (99.9)	999	99	999	999	999.9	99.9	999	99 (99.9)	999		99	999	999	999.9	99.9	999	99 (99.9)	999

NOTE: Use variable HGT_MOM

28. Previous preterm infant - <37 wks
    For each Data Partner and Totals
    (Percents are column percents within State)

	California		Colorado		Florida		*etc.*		Total	
	N	%	N	%	N	%	N	%	N	%
Yes	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
No	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0

NOTE: Use variable PRV_LT37

29. Previous infant small for gest age (SGA) or previous preterm infant
    For each Data Partner and Totals
    (Percents are column percents within State)

	California		Colorado		Florida		*etc.*		Total	
	N	%	N	%	N	%	N	%	N	%
Yes	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
No	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0

NOTE: Use variable PRVSMALL_LT37

30. No congenital anomalies listed By Year
    For each Data Partner and Totals
    (Percents are column percents within State/Year)

		California		Colorado		Florida		*etc.*		Total	
Birth Year		N	%	N	%	N	%	N	%	N	%
2007	Yes	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	No	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0
2008	Yes	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	No	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0
Etc.	Yes	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	No	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9

		California		Colorado		Florida		*etc.*		Total	
Birth Year		**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
	Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0
Total	Yes	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	No	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Missing	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9	99,999	99.9
	Totals	99,999	100.0	99,999	100.0	99,999	100.0	99,999	99.9	99,999	100.0

NOTE: Use variable NO_CONG_AN

31. Spina bifida/meningocele

For each Data Partner and Totals
(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable BIFIDA

32. Anencephalus

For each Data Partner and Totals
(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable ANENCEPH

33. Heart malformations

For each Data Partner and Totals
(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable MALF_HRT

34. Cyanotic congenital heart disease

For each Data Partner and Totals
(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable CYAN_CONG_HEART

35. Omphalocele

For each Data Partner and Totals
(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable OMPHALO

36. Gastroschisis

For each Data Partner and Totals
(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable GASTROSCH

37. Unspecified if omphalocele or gastroschisis

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable OMPHALO_UNSP

38. Tra/esophageal fistula, atresia

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable TRACH_ESO_FISTUL

39. Hypospadias

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable HYPOSPAD

40. Cleft lip with or without cleft palate

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable C_LIP

41. Cleft palate alone

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable PALATE_ONLY

42. Cleft lip/palate unspecified

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable PALATE_UNSP

43. Limb reduction defect

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable LIMB_RED

44. Diaphragmatic hernia

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable DIAPH_HERNIA

45. Downs syndrome

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable DOWNS

46. Other chromosomal anomalies

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable OTHERCHR

47. Diabetes

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable DIABETES

48. Gestational diabetes

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable DIAB_GEST

49. Nongestational (preexisting) diabetes

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable DIAB_NONGEST

50. Chronic hypertension

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable CHYPER

51. Hypertension/preeclampsia

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable PIH

52. Eclampsia

For each Data Partner and Totals

(Percents are column percents within State/Year)

Table follows format of Number 30. No congenital anomalies listed By Year

NOTE: Use variable ECLAMP

C. APPENDIX C. PROPOSED MINI-SENTINEL COMMON DATA MODEL (MSCDM) BIRTH AND FETAL DEATH TABLES

MSCDM Data File Specifications: Birth Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
GENERAL BIRTH INFORMATION				
MPatID	Char(##)	$##.	Mother's patient ID	Text string; left justified
CPatID	Char(##)	$##.	Child's patient ID	Text string; left justified
STATE	Char(3)	$3.	State	State postal abbreviations, with the addition of "NYC"
BDOB	N(4)	mmddyy10.	Child's Date of Birth	In years 20## - 2012
BSEX	Char(1)	$1.	Child sex	M = Male F = Female A = Ambiguous (e.g., transgender/hermaphrodite) U = Unknown
GESTCLIN	Num(3)	2.	Gestational age clinical estimate (wks)	In weeks . = not reported, missing
GESTMENS	Num(3)	2.	Gestational age based on last menses (wks)	In weeks . = not reported, missing
GESTOBSTET	Num(3)	2.	Gestational age based on obstetric estimate	In weeks . = not reported, missing
DAYSGEST_OTH	Num(3)	3.	Days of Gestation	In days . = not reported, missing
HOW	Num(3)	2.	Method specified for DAYSGEST_OTH	1 data provided in file 2 imputed from birthweight 3 other -1 basis for days of gestation is unknown . Missing

MSCDM Data File Specifications: Birth Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
DELIVMETH	Char(2)	$2.	Delivery method	VA = vaginal VB = VBAC FO = Forceps VC = Vacuum PC = Primary C-section RC= Repeat C-section CU= C-section unspecified OT= Other UK = unknown or unreported
BWEIGHT	Num (4)	4.	Weight of infant in grams	Positive integers for birthweight in grams . Missing
PLURALITY	Num(3)	1.	Single, twin, etc	positive integers 1-8: 1=single 2=twin 3=triplet 4=quadruplet 5=quintuplet 6=sextuplet 7=septuplet 8=8 or more or Missing (unknown/unreported)
MOTHER AND FATHER INFORMATION				
MDOB	Num(4)	mmddyy10.	Mother's Date of Birth	In years 19## - 20## or missing
MMARSTATUS	Char(1)	$1.	Mother's marital status	M = Married S = Single D = Divorced W = Widowed O = Other N= Not married (unspecified if Single, Divorced, or Widowed) U = Unknown
MENSDT	Num(4)	mmddyy10.	Mother's date of last menses	In years 19## - 20## or missing

MSCDM Data File Specifications: Birth Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
PRENATMON	Num(3)	3.	Month prenatal care began	0 = No prenatal care 1 = First Month 2 = Second Month 3 = Third Month 4 = Fourth Month 5 = Fifth Month 6 = Sixth Month 7 = Seventh Month 8 = Eighth Month 9 = Ninth month or later or Missing/Unknown
PRENATNUM	Num(3)	2.	Number of prenatal care visits	0 = No prenatal care or positive integers or Missing
MGRAVIDITY	Num(3)	3.	Gravidity (total pregnancies)	Missing, zero or positive integers
NUM_BTHS	Num(3)	3.	Previous live births total	Missing, zero or positive integers
PLIV_DEAD	Num(3)	3.	Previous live births now dead	Missing, zero or positive integers
PLIV_LIV	Num(3)	3.	Previous live births now living	Missing, zero or positive integers
NUM_TRMS	Num(3)	3.	Previous other pregnancy outcomes	Missing, zero or positive integers

MSCDM Data File Specifications: Birth Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
MRACE	Char(2)	$2.	Mother's race	Two-digit SEER code. '01' = "White" '02' = "Black" '03' = "American Indian, Aleutian, or Eskimo" '04' = "Chinese" '05' = "Japanese" '06' = "Filipino" '07' = "Hawaiian" '08' = "Korean" '09' = "Asian Indian, Pakistani" '10' = "Vietnamese" '11' = "Laotian" '12' = " Hmong " '13' = "Kampuchean" '14' = "Thai" '20' = "Micronesian, NOS" '21' = " Chamorran " '22' = " Guamanian , NOS" '25' = "Polynesian, NOS" '26' = "Tahitian" '27' = "Samoan" '28' = "Tongan" '30' = "Melanesian, NOS" '31' = "Fiji Islander" '32' = "New Guinean" '96' = "Other Asian, incl. Asian, NOS and Oriental, NOS"--also used when multiple Asian races recorded, or Asian race not specifically categorized as above '97' = "Pacific Islander, NOS"-- also used when multiple Pacific Islander races recorded, or Pacific Island race not specifically categorized as above '98' = "Other"-- also used if race not specifically categorized as above '99' = "Unknown" 'MU' = Multi-racial

MSCDM Data File Specifications: Birth Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
FRACE	Char(2)	$2.	Father's race	Two-digit SEER code. '01' = "White" '02' = "Black" '03' = "American Indian, Aleutian, or Eskimo" '04' = "Chinese" '05' = "Japanese" '06' = "Filipino" '07' = "Hawaiian" '08' = "Korean" '09' = "Asian Indian, Pakistani" '10' = "Vietnamese" '11' = "Laotian" '12' = " Hmong " '13' = "Kampuchean" '14' = "Thai" '20' = "Micronesian, NOS" '21' = " Chamorran " '22' = " Guamanian , NOS" '25' = "Polynesian, NOS" '26' = "Tahitian" '27' = "Samoan" '28' = "Tongan" '30' = "Melanesian, NOS" '31' = "Fiji Islander" '32' = "New Guinean" '96' = "Other Asian, incl. Asian, NOS and Oriental, NOS"--also used when multiple Asian races recorded, or Asian race not specifically categorized as above '97' = "Pacific Islander, NOS"-- also used when multiple Pacific Islander races recorded, or Pacific Island race not specifically categorized as above '98' = "Other"-- also used if race not specifically categorized as above '99' = "Unknown" 'MU' = Multi-racial

MSCDM Data File Specifications: Birth Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
HISP_MOM	Num(3)	2.	Hispanic origin - mother	. missing 0 Non-Hispanic or No 1 Mexican 2 Puerto Rican 3 Cuban 4 Central or South America, Spanish speaking countries 5 Other & Unknown Hispanic
HISP_DAD	Num(3)	2.	Hispanic origin - father	. missing 0 Non-Hispanic or No 1 Mexican 2 Puerto Rican 3 Cuban 4 Central or South America, Spanish speaking countries 5 Other & Unknown Hispanic
AGE_MOM	Num(3)	2.	Age of mother at birth of child	Positive integers or Missing
AGE_DAD	Num(3)	2.	Age of father at birth of child	Positive integers or Missing
EDUC_MOM	Char(2)	$2.	Education of mother	00 = No Education 08 = 8th grade or less 11 = 9th grade through 12th grade, no diploma 12 = High school graduate or GED completed 13 = Some college credit but no degree 14 = Associate's Degree 16 = Bachelor's Degree 17 = Education beyond Bachelor's Degree 88 = One or more years of college, not stated whether degree earned 99 = missing; unknown

MSCDM Data File Specifications: Birth Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
EDUC_DAD	Char(2)	$2.	education of father	00 = No Education 08 = 8th grade or less 11 = 9th grade through 12th grade, no diploma 12 = High school graduate or GED completed 13 = Some college credit but no degree 14 = Associate's Degree 16 = Bachelor's Degree 17 = Education beyond Bachelor's Degree 88 = One or more years of college, not stated whether degree earned 99 = missing; unknown
TOBACCO	Num(3)	1.	mother was a smoker	1 Yes 0 No Missing
MCIGNUM	Num(3)	3.	Number of cigarettes daily during preg	0, positive integer or -88 less than one, occasional use -89 chewing tobacco or Missing
CIGPREPREG	Num(3)	3.	Number of cigarettes daily pre-pregnancy	0, positive integer or -88 less than one, occasional use -89 chewing tobacco or Missing
CIGTRIM1	Num(3)	3.	Number of cigarettes daily during 1st 3 months of pregnancy	0, positive integer or -88 less than one, occasional use -89 chewing tobacco or Missing
CIGTRIM2	Num(3)	3.	Number of cigarettes daily during 2nd 3 months of pregnancy	0, positive integer or -88 less than one, occasional use -89 chewing tobacco or Missing
CIGTRIM3	Num(3)	3.	Number of cigarettes daily during 3rd trimester	0, positive integer or -88 less than one, occasional use -89 chewing tobacco or Missing

MSCDM Data File Specifications: Birth Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
ALCOHOL	Num(3)	1.	Mother drank alcohol	1 Yes 0 No Missing
DRINKS	Num(3)	3.	Number drinks consumed per week	0, positive integer or -88 less than one, occasional use or Missing
WGT_PRE_PREG	Num(3)	3.	Maternal weight pre-pregnancy	Positive integers or Missing
HGT_MOM	Num(4)	3.2	Maternal height in meters	Positive number or Missing
PRV_LT37	Num(3)	1.	Previous preterm infant - <37 wks	1 Yes 2 No Missing
PRVSMALL_LT37	Num(3)	1.	Previous infant small for gest age (SGA) or previous preterm infant	1 Yes 2 No Missing
ABNORMAL NEWBORN CONDITIONS, CONGENITAL ANOMALIES				
NO_CONG_AN	Num(3)	1.	No congenital anomalies listed	1 Yes 2 No Missing
BIFIDA	Num(3)	1.	Spina bifida/meningocele	1 Yes 2 No Missing
ANENCEPH	Num(3)	1.	Anencephalus	1 Yes 2 No Missing
MALF_HRT	Num(3)	1.	Heart malformations	1 Yes 2 No Missing
CYAN_CONG_HEART	Num(3)	1.	Cyanotic congenital heart disease	1 Yes 2 No Missing
OMPHALO	Num(3)	1.	Omphalocele	1 Yes 2 No Missing

MSCDM Data File Specifications: Birth Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
GASTROSCH	Num(3)	1.	Gastroschisis	1 Yes 2 No Missing
OMPHALO_UNSP	Num(3)	1.	Unspecified if omphalocele or gastroschisis	1 Yes 2 No Missing
TRACH_ESO_FISTUL	Num(3)	1.	Tra/esophageal fistula, atresia	1 Yes 2 No Missing
HYPOSPAD	Num(3)	1.	Hypospadias	1 Yes 2 No Missing
C_LIP	Num(3)	1.	Cleft lip with or without cleft palate	1 Yes 2 No Missing
PALATE_ONLY	Num(3)	1.	Cleft palate alone	1 Yes 2 No Missing
PALATE _UNSP	Num(3)	1.	Cleft lip/palate unspecified	1 Yes 2 No Missing
LIMB_RED	Num(3)	1.	Limb reduction defect	1 Yes 2 No Missing
DIAPH_HERNIA	Num(3)	1.	Diaphragmatic hernia	1 Yes 2 No Missing
DOWNS	Num(3)	1.	Downs syndrome	1 Yes 2 No 3 Pending Missing
OTHERCHR	Num(3)	1.	Other chromosomal anomalies	1 Yes 2 No 3 Pending Missing
COMPLICATIONS OF PREGNANCY AND CONCURRENT ILLNESSES				

MSCDM Data File Specifications: Birth Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
DIABETES	Num(3)	1.	Diabetes	1 Yes 2 No Missing
DIAB_GEST	Num(3)	1.	Gestational diabetes	1 Yes 2 No Missing
DIAB_NONGEST	Num(3)	1.	Nongestational (preexisting) diabetes	1 Yes 2 No Missing
CHYPER	Num(3)	1.	Chronic hypertension	1 Yes 2 No Missing
PIH	Num(3)	1.	Hypertension/preeclampsia	1 Yes 2 No Missing
ECLAMP	Num(3)	1.	Eclampsia	1 Yes 2 No Missing

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
GENERAL BIRTH INFORMATION				
MPatID	Char(##)	$##.	Mother's patient ID	Text string; left justified
STATE	Char(3)	$3.	State	State postal abbreviations, with the addition of "NYC"
FDDate	Num(4)	mmddyy10.	Date of fetal delivery	In years 20## - 2012
Fsex	Char(1)	$1.	Sex of fetus	M = Male F = Female A = Ambiguous (e.g., transgender/hermaphrodite) U = Unknown
GESTCLIN	Num(3)	2.	Gestational age clinical estimate (wks)	In weeks . = not reported, missing
GESTMENS	Num(3)	2.	Gestational age based on last menses (wks)	In weeks . = not reported, missing
GESTOBSTET	Num(3)	2.	Gestational age based on obstetric estimate	In weeks . = not reported, missing
DAYSGEST_OTH	Num(3)	3.	Days of Gestation	In days . = not reported, missing
HOW	Num(3)	2.	Method specified for DAYSGEST_OTH	1 data provided in file 2 imputed from birthweight 3 other -1 basis for days of gestation is unknown . Missing
BWEIGHT	Num (4)	4.	Weight of infant in grams	Positive integers for birthweight in grams . Missing

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
PLURALITY	Num(3)	1.	Single, twin, etc	positive integers 1-8: 1=single 2=twin 3=triplet 4=quadruplet 5=quintuplet 6=sextuplet 7=septuplet 8=8 or more or Missing (unknown/unreported)
MOTHER AND FATHER INFORMATION				
MDOB	Num(4)	mmddyy10.	Mother's Date of Birth	In years 19## - 20## or missing
MMARSTATUS	Char(1)	$1.	Mother's marital status	M = Married S = Single D = Divorced W = Widowed O = Other N= Not married (unspecified if Single, Divorced, or Widowed) U = Unknown
MENSDT	Num(4)	mmddyy10.	Mother's date of last menses	In years 19## - 20## or missing
PRENATMON	Num(3)	3.	Month prenatal care began	0 = No prenatal care 1 = First Month 2 = Second Month 3 = Third Month 4 = Fourth Month 5 = Fifth Month 6 = Sixth Month 7 = Seventh Month 8 = Eighth Month 9 = Ninth month or later or Missing/Unknown

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
PRENATNUM	Num(3)	2.	Number of prenatal care visits	0 = No prenatal care or positive integers or Missing
MGRAVIDITY	Num(3)	3.	Gravidity (total pregnancies)	Missing, zero or positive integers
NUM_BTHS	Num(3)	3.	Previous live births total	Missing, zero or positive integers
PLIV_DEAD	Num(3)	3.	Previous live births now dead	Missing, zero or positive integers
PLIV_LIV	Num(3)	3.	Previous live births now living	Missing, zero or positive integers
NUM_TRMS	Num(3)	3.	Previous other pregnancy outcomes	Missing, zero or positive integers

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
MRACE	Char(2)	$2.	Mother's race	Two-digit SEER code. '01' = "White" '02' = "Black" '03' = "American Indian, Aleutian, or Eskimo" '04' = "Chinese" '05' = "Japanese" '06' = "Filipino" '07' = "Hawaiian" '08' = "Korean" '09' = "Asian Indian, Pakistani" '10' = "Vietnamese" '11' = "Laotian" '12' = " Hmong " '13' = "Kampuchean" '14' = "Thai" '20' = "Micronesian, NOS" '21' = " Chamorran " '22' = " Guamanian , NOS" '25' = "Polynesian, NOS" '26' = "Tahitian" '27' = "Samoan" '28' = "Tongan" '30' = "Melanesian, NOS" '31' = "Fiji Islander" '32' = "New Guinean" '96' = "Other Asian, incl. Asian, NOS and Oriental, NOS"--also used when multiple Asian races recorded, or Asian race not specifically categorized as above '97' = "Pacific Islander, NOS"-- also used when multiple Pacific Islander races recorded, or Pacific Island race not specifically categorized as above '98' = "Other"-- also used if race not specifically categorized as above '99' = "Unknown" 'MU' = Multi-racial

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
FRACE	Char(2)	$2.	Father's race	Two-digit SEER code. '01' = "White" '02' = "Black" '03' = "American Indian, Aleutian, or Eskimo" '04' = "Chinese" '05' = "Japanese" '06' = "Filipino" '07' = "Hawaiian" '08' = "Korean" '09' = "Asian Indian, Pakistani" '10' = "Vietnamese" '11' = "Laotian" '12' = " Hmong " '13' = "Kampuchean" '14' = "Thai" '20' = "Micronesian, NOS" '21' = " Chamorran " '22' = " Guamanian , NOS" '25' = "Polynesian, NOS" '26' = "Tahitian" '27' = "Samoan" '28' = "Tongan" '30' = "Melanesian, NOS" '31' = "Fiji Islander" '32' = "New Guinean" '96' = "Other Asian, incl. Asian, NOS and Oriental, NOS"--also used when multiple Asian races recorded, or Asian race not specifically categorized as above '97' = "Pacific Islander, NOS"-- also used when multiple Pacific Islander races recorded, or Pacific Island race not specifically categorized as above '98' = "Other"-- also used if race not specifically categorized as above '99' = "Unknown" 'MU' = Multi-racial

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
HISP_MOM	Num(3)	2.	Hispanic origin - mother	. missing 0 Non-Hispanic or No 1 Mexican 2 Puerto Rican 3 Cuban 4 Central or South America, Spanish speaking countries 5 Other & Unknown Hispanic
HISP_DAD	Num(3)	2.	Hispanic origin - father	. missing 0 Non-Hispanic or No 1 Mexican 2 Puerto Rican 3 Cuban 4 Central or South America, Spanish speaking countries 5 Other & Unknown Hispanic
AGE_MOM	Num(3)	2.	age of mother at delivery of fetus	Positive integers or Missing
AGE_DAD	Num(3)	2.	age of father at delivery of fetus	Positive integers or Missing
EDUC_MOM	Char(2)	$2.	Education of mother	00 = No Education 08 = 8th grade or less 11 = 9th grade through 12th grade, no diploma 12 = High school graduate or GED completed 13 = Some college credit but no degree 14 = Associate's Degree 16 = Bachelor's Degree 17 = Education beyond Bachelor's Degree 88 = One or more years of college, not stated whether degree earned 99 = missing; unknown

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
EDUC_DAD	Char(2)	$2.	Education of father	00 = No Education 08 = 8th grade or less 11 = 9th grade through 12th grade, no diploma 12 = High school graduate or GED completed 13 = Some college credit but no degree 14 = Associate's Degree 16 = Bachelor's Degree 17 = Education beyond Bachelor's Degree 88 = One or more years of college, not stated whether degree earned 99 = missing; unknown
TOBACCO	Num(3)	1.	Mother was a smoker	1 Yes 0 No Missing
MCIGNUM	Num(3)	3.	Number of cigarettes daily during preg	0, positive integer or -88 less than one, occasional use -89 chewing tobacco or Missing
CIGPREPREG	Num(3)	3.	Number of cigarettes daily pre-pregnancy	0, positive integer or -88 less than one, occasional use -89 chewing tobacco or Missing
CIGTRIM1	Num(3)	3.	Number of cigarettes daily during 1st 3 months of pregnancy	0, positive integer or -88 less than one, occasional use -89 chewing tobacco or Missing
CIGTRIM2	Num(3)	3.	Number of cigarettes daily during 2nd 3 months of pregnancy	0, positive integer or -88 less than one, occasional use -89 chewing tobacco or Missing
CIGTRIM3	Num(3)	3.	Number of cigarettes daily during 3rd trimester	0, positive integer or -88 less than one, occasional use -89 chewing tobacco or Missing

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
ALCOHOL	Num(3)	1.	Mother drank alcohol	1 Yes 0 No Missing
DRINKS	Num(3)	3.	Number drinks consumed per week	0, positive integer or -88 less than one, occasional use or Missing
WGT_PRE_PREG	Num(3)	3.	Maternal weight pre-pregnancy	Positive integers or Missing
HGT_MOM	Num(4)	3.2	Maternal height in meters	Positive integers or Missing
PRV_LT37	Num(3)	1.	Previous preterm infant - <37 wks	1 Yes 2 No Missing
PRVSMALL_LT37	Num(3)	1.	Previous infant small for gest age (SGA) or previous preterm infant	1 Yes 2 No Missing
ABNORMAL NEWBORN CONDITIONS, CONGENITAL ANOMALIES				
NO_CONG_AN	Num(3)	1.	No congenital anomalies listed	1 Yes 2 No Missing
BIFIDA	Num(3)	1.	Spina bifida/meningocele	1 Yes 2 No Missing
ANENCEPH	Num(3)	1.	Anencephalus	1 Yes 2 No Missing
MALF_HRT	Num(3)	1.	Heart malformations	1 Yes 2 No Missing
CYAN_CONG_HEART	Num(3)	1.	Cyanotic congenital heart disease	1 Yes 2 No Missing
OMPHALO	Num(3)	1.	Omphalocele	1 Yes 2 No Missing

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
GASTROSCH	Num(3)	1.	Gastroschisis	1 Yes 2 No Missing
OMPHALO_UNSP	Num(3)	1.	Unspecified if omphalocele or gastroschisis	1 Yes 2 No Missing
TRACH_ESO_FISTUL	Num(3)	1.	Tra/esophageal fistula, atresia	1 Yes 2 No Missing
HYPOSPAD	Num(3)	1.	Hypospadias	1 Yes 2 No Missing
C_LIP	Num(3)	1.	Cleft lip with or without cleft palate	1 Yes 2 No Missing
PALATE_ONLY	Num(3)	1.	Cleft palate alone	1 Yes 2 No Missing
PALATE _UNSP	Num(3)	1.	Cleft lip/palate unspecified	1 Yes 2 No Missing
LIMB_RED	Num(3)	1.	Limb reduction defect	1 Yes 2 No Missing
DIAPH_HERNIA	Num(3)	1.	Diaphragmatic hernia	1 Yes 2 No Missing
DOWNS	Num(3)	1.	Downs syndrome	1 Yes 2 No 3 Pending Missing
OTHERCHR	Num(3)	1.	Other chromosomal anomalies	1 Yes 2 No 3 Pending Missing
COMPLICATIONS OF PREGNANCY AND CONCURRENT ILLNESSES				

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
DIABETES	Num(3)	1.	Diabetes	1 Yes 2 No Missing
DIAB_GEST	Num(3)	1.	Gestational diabetes	1 Yes 2 No Missing
DIAB_NONGEST	Num(3)	1.	Nongestational (preexisting) diabetes	1 Yes 2 No Missing
CHYPER	Num(3)	1.	Chronic hypertension	1 Yes 2 No Missing
PIH	Num(3)	1.	Hypertension/preeclampsia	1 Yes 2 No Missing
ECLAMP	Num(3)	1.	Eclampsia	1 Yes 2 No Missing
CAUSE OF DEATH				
COD_INT1_ICD10	Char (8)	$8.	Immediate/initiating cause of death (ICD-10)	Text string; left justified
COD_INT2_ICD10	Char (8)	$8.	Immediate/initiating cause of death (ICD-10)	Text string; left justified
COD_INT_MATERNAL	Char(1)	$1.	immediate/initiating cause - maternal	Y = Yes N = No U = Unknown/missing
COD_INT_FETAL	Char(1)	$1.	immediate/initiating cause - fetal	Y = Yes N = No U = Unknown/missing
COD_INT_DUETO_1_MF	Char(1)	$1.	immediate/initiating cause due to condition 1- maternal vs fetal	Y = Yes N = No U = Unknown/missing
COD_INT_DUETO_2_MF	Char(1)	$1.	immediate/initiating cause due to condition 2- maternal vs fetal	Y = Yes N = No U = Unknown/missing

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
COD_INT_COMP_PLAC	Char(1)	$1.	complications of placenta, cord, membranes - immediate/initiating cause	Y = Yes N = No U = Unknown/missing
COD_INT_RUPT_MEM	Char(1)	$1.	rupture of membranes - immediate/initiating cause	Y = Yes N = No U = Unknown/missing
COD_INT_ABRUPTIO	Char(1)	$1.	abruptio placenta - immediate/initiating cause	Y = Yes N = No U = Unknown/missing
COD_INT_PLAC_INSUFF	Char(1)	$1.	placental insufficiency - immediate/initiating cause	Y = Yes N = No U = Unknown/missing
COD_INT_PROLAP_CORD	Char(1)	$1.	prolapsed cord - immediate/initiating cause	Y = Yes N = No U = Unknown/missing
COD_INT_TRU_KNOT	Char(1)	$1.	true knot in cord- immediate/initiating cause	Y = Yes N = No U = Unknown/missing
COD_INT_CHORIO	Char(1)	$1.	chorioamnionitis - immediate/initiating cause	Y = Yes N = No U = Unknown/missing
COD_INT_OTH_OB	Char(1)	$1.	other obstetrical complic - immediate/initiating cause	Y = Yes N = No U = Unknown/missing
COD_INT_FETAL_ANOM	Char(1)	$1.	fetal anomaly - immediate/initiating cause	Y = Yes N = No U = Unknown/missing
COD_INT_FETAL_INJ	Char(1)	$1.	fetal injury - immediate/initiating cause	Y = Yes N = No U = Unknown/missing
COD_INT_FETAL_INF	Char(1)	$1.	fetal infection - immediate/initiating cause	Y = Yes N = No U = Unknown/missing
COD_INT_OTH_FETAL	Char(1)	$1.	other fetal condition- immediate/initiating cause	Y = Yes N = No U = Unknown/missing

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
COD_INT_CAUSE_UNK	Char(1)	$1.	unknown immediate cause	Y = Yes N = No U = Unknown/missing
COD_UNDERLYING_ICD10	Char (8)	$8.	Underlying cause of death (ICD-10)	Text string; left justified
COD_OTH_ICD_CODE1	Char (8)	$8.	other significant cause of death (ICD-10)	Text string; left justified
COD_OTH_MAT1	Char(1)	$1.	other significant cause - maternal	Y = Yes N = No U = Unknown/missing
COD_OTH_FETAL1	Char(1)	$1.	other significant cause - fetal	Y = Yes N = No U = Unknown/missing
COD_OTH_ICD_CODE2	Char (8)	$8.	other significant cause of death (ICD-10)	Text string; left justified
COD_OTH_MAT2	Char(1)	$1.	other significant cause - maternal	Y = Yes N = No U = Unknown/missing
COD_OTH_FETAL2	Char(1)	$1.	other significant cause - fetal	Y = Yes N = No U = Unknown/missing
COD_OTH_ICD_CODE3	Char (8)	$8.	other significant cause of death (ICD-10)	Text string; left justified
COD_OTH_ICD_CODE4	Char (8)	$8.	other significant cause of death (ICD-10)	Text string; left justified
COD_OTH_ICD_CODE5	Char (8)	$8.	other significant cause of death (ICD-10)	Text string; left justified
COD_OTH_ICD_CODE6	Char (8)	$8.	other significant cause of death (ICD-10)	Text string; left justified
COD_OTH_ICD_CODE7	Char (8)	$8.	other significant cause of death (ICD-10)	Text string; left justified
COD_OTH_COMP_PLAC	Char(1)	$1.	complications of placenta, cord, membranes - other significant cause	Y = Yes N = No U = Unknown/missing

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
COD_OTH_RUPT_MEM	Char(1)	$1.	rupture of membranes - other significant cause	Y = Yes N = No U = Unknown/missing
COD_OTH_ABRUPTIO	Char(1)	$1.	abruptio placenta - other significant cause	Y = Yes N = No U = Unknown/missing
COD_OTH_PLAC_INSUFF	Char(1)	$1.	placental insufficiency - other significant cause	Y = Yes N = No U = Unknown/missing
COD_ OTH_PROLAP_CORD	Char(1)	$1.	prolapsed cord - other significant cause	Y = Yes N = No U = Unknown/missing
COD_OTH_TRU_KNOT	Char(1)	$1.	true knot in cord- other significant cause	Y = Yes N = No U = Unknown/missing
COD_OTH_CHORIO	Char(1)	$1.	chorioamnionitis - other significant cause	Y = Yes N = No U = Unknown/missing
COD_OTH_OTH_OB	Char(1)	$1.	other obstetrical complic - other significant cause	Y = Yes N = No U = Unknown/missing
COD_OTH_FETAL_ANOM	Char(1)	$1.	fetal anomaly - other significant cause	Y = Yes N = No U = Unknown/missing
COD_OTH_FETAL_INJ	Char(1)	$1.	fetal injury - other significant cause	Y = Yes N = No U = Unknown/missing
COD_OTH_FETAL_INF	Char(1)	$1.	fetal infection - other significant cause	Y = Yes N = No U = Unknown/missing
COD_OTH_FETAL	Char(1)	$1.	other fetal condition- other significant cause	Y = Yes N = No U = Unknown/missing
OTH_CAUSE_UNK	Char(1)	$1.	unknown other significant cause	Y = Yes N = No U = Unknown/missing

MSCDM Data File Specifications: Fetal Death Certificate Table				
Variable Name	**Data Type**	**Format**	**Label**	**Valid Values**
AUTOPSY	Char(1)	$1.	autopsy performed	Y = Yes N = No U = Unknown/missing
AUTOPSY_DETERM	Char(1)	$1.	autopsy or histologic exam used in the determination of the cause of death	Y = Yes N = No U = Unknown/missing

D. APPENDIX D. FDA LETTER TO NEW YORK CITY DEPARTMENT OF HEALTH AND MENTAL HYGIENE (DOHMH)

DEPARTMENT OF HEALTH & HUMAN SERVICES

Food and Drug Administration
10903 New Hampshire Avenue
Building #51
Silver Spring, MD 20993

October 08, 2013

Dear Dr. Farley,

The U.S. Food and Drug Administration (FDA) is developing a new vaccine safety monitoring program for pregnant women within Mini-Sentinel's Post Licensure Rapid Immunization Safety Monitoring (PRISM) program. The Mini-Sentinel Operations Center (MSOC) and two of our collaborating institutions (HealthCore and Aetna) have recently been in communication with your Office of Vital Statistics to request the matching of their members to birth registry data to augment PRISM's Pregnancy Safety Surveillance Program. This request was denied by your office on the grounds that the Data Health Code does not permit disclosure to a non-government entity.

I am writing to request that you reconsider this decision. As explained in the accompanying letter, which was included with the initial application, the Mini-Sentinel Operations Center and Collaborating Institutions are also considered public health authorities for purposes of the Mini-Sentinel pilot, because they are acting under contract with and under the authority of the FDA.

Mini Sentinel would be happy to arrange for their representative to meet with your legal team to clarify any questions they may have about the status of our collaborating institutions.

FDA is dedicated to ensuring the safety of vaccines in pregnancy and is working with the PRISM program to develop and test an advanced surveillance system. The supplemental data which would be gained by incorporating birth certificate data into the Mini-Sentinel model would be invaluable in improving FDA's ability to monitor for adverse pregnancy outcomes such as miscarriages, premature births, and birth defects.

We hope that we can resolve this issue in a timely manner so that New York City Vital Statistics can become a major contributor to this important public health activity.

Please contact Carlos.Bell@fda.hhs.gov , FDA's Sentinel Program Manager, if you have any questions or concerns about this letter, or the PRISM Project Manager, Linda Pointon (linda_pointon@harvardpilgrim.org) for further information about the safety assessment.

Sincerely,

Rachel E. Sherman, MD, MPH, FACP
Associate Director of Medical Policy
Center for Drug Evaluation and Research
US Food and Drug Administration