Maternal Morbidity and Outcomes Including Mortality, California 2001-2006
Appendix A: Methods

Linda L. Remy, MSW PhD
Ted Clay, MS
Geraldine Oliva, MD MPH

April 2008

Family Health Outcomes Project
University of California, San Francisco
www.ucsf.edu/FHOP

Funding Provided by a Grant to the Bay Area Data Collaborative
from the Maternal Care Quality Improvement Project of the
California Department of Public Health, Maternal Child and Adolescent Health Branch.

APPENDIX A: METHODS

Goals and Questions

- Identify regional variations in maternal morbidity and mortality
    - Associated with poor outcomes
    - Large population impact
    - Amenable to intervention
- Prepare health jurisdiction databooks highlighting important conditions to monitor

Given these goals, our basic questions were as follows:

1. Between 2001 and 2006, how many pregnant women age 10 to 60 sought care in California hospitals? A delivery outcomes study done as part of OSHPD's California Hospital Outcomes Project used the age range 10 to 55 years,[1] while Hornbrook used the range 12 to 60 to study maternal outcomes.[2] Note that these age frames are wider than the range 15 to 44 traditionally used to study maternal indicators for public health surveillance. We broadened the age frame to address a related question in the California death certificates. Specifically, if the deceased woman is between age 10 and 60, the person completing the death certificate is supposed to report if she was pregnant in the last year.

2. How many women had an adverse pregnancy outcome or died within 1 year of a pregnancy-related discharge? This attempts to answer the question on the death certificate. In essence, the first question gives us a denominator. The second question gives numerators for patient safety indicators including mortality.

3. What differentiated women who had adverse outcomes or died? To answer this question, we compared differences in diagnoses (including comorbid conditions), and procedures for women who had adverse outcomes including death.

4. Do differences suggest interventions? From our analysis, we identified several possible interventions for the Bay Area Data Collaborative to pursue.

Data Sources

To address these questions, we turned to California's wealth of administrative datasets. These data offer a window into medical care delivered in our nation’s hospitals, emergency departments and clinics. They are collected as a routine step in the delivery of care from birth to death. They provide information on diagnoses, procedures, age, gender, admission source, and discharge status. From these data, it is possible to construct a picture of medical care quality. Although quality assessments based on administrative data cannot be definitive, they can be used to flag potential quality problems and successes, which can then be further investigated and studied.[3] All work was done using SAS 9.1.3 Service Pack 4.

Office of Statewide Health Planning and Development (OSHPD)

For the period 2001 to 2006 we used Patient Discharge Data (PDD) from the Office of Statewide Health Planning and Development (OSHPD). For 2005-2006, we used OSHPD's newly available Emergency Department (ED) and Ambulatory Care (ambulatory surgery or clinic (AC)) datasets. Many pregnancy-related discharges, including deliveries, were found in ED/AS datasets. Thus pregnancy-related discharges were not limited to the inpatient setting. We use the term "discharge" to cover inpatient and outpatient events in all settings OSHPD licenses.

In each OSHPD dataset, we pulled all California resident women whose age at admission was between 10 and 60 to be consistent with the parallel question in the Vital Statistics Death Certificate: Was this woman pregnant in the last year?. We identified residence by ZIP-codes in the California range 90000 and 96162. We initially also kept all non-pregnancy deaths in this range, presuming that some number of those would turn out to be deaths of women who had been pregnant. Finally, after we summarized the data to a "woman" we returned to these datasets and pulled all other non-pregnancy discharges.

Vital Statistics

Our initial plan had been to use California's Vital Statistics death files for 2001-2006. Vital Statistics still has not released the 2006 deaths, so this analysis used deaths through 2005. All deaths within one year after delivery are not available at this time for women pregnant in 2005 or 2006. As a result, deaths are undercounted.

We made an initial file containing all deaths of women age 10 to 60, consistent with the pregnancy question in the death certificate. These were less than 10% of all California deaths. The death files use the International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM) mortality codes to describe the underlying cause of death.

Event Selection and Classification

Identify and Classify Events for Pregnant Women

OSHPD datasets list the patient’s birth date, sex, race, ZIP code, expected source of payment, disposition, social security number (SSN), principal and up to 24 secondary diagnoses, up to 21 procedures, and up to 5 external causes of injury. The PDD also includes admission source and type of admission, total charges, Major Diagnostic Group (MDC) and Diagnosis Related Group (DRG). All datasets describe diagnoses using International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes. The PDD use ICD-9-CM procedure codes, while the ED/AS have Current Procedural Terminology (CPT) codes. CPT is a proprietary coding system developed by the American Medical Association.

To identify pregnancy-related and other conditions, we used software developed by the federal Agency for Healthcare Research and Quality. These software allow uniform calculation of indicators to measure progress toward the nation's Healthy People 2010 goals and objectives.

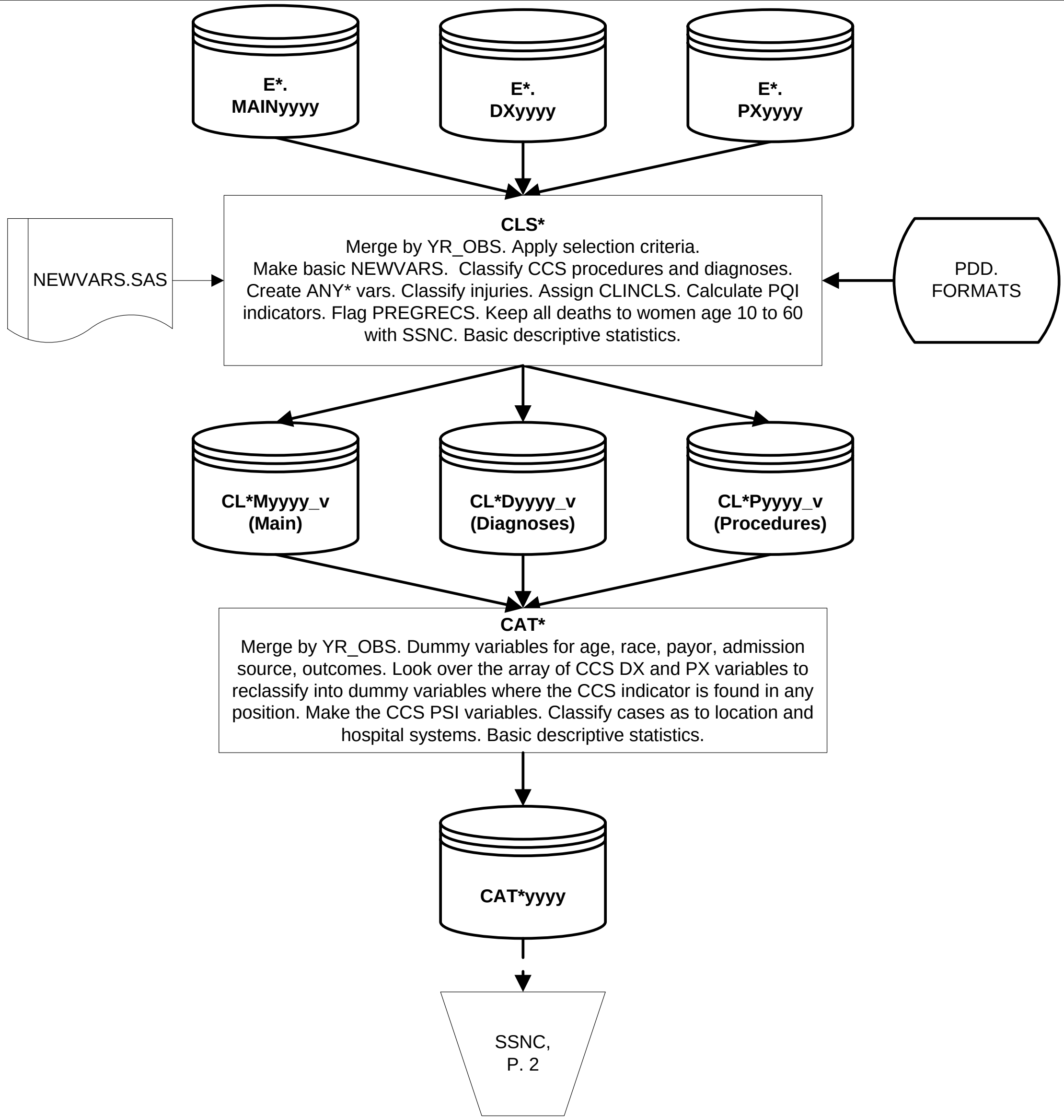
- **Clinical Classification System** (CCS) is a tool for clustering patient diagnoses and procedures into a manageable number of clinically meaningful categories.[4] CCS offers researchers the ability to group conditions and procedures without having to sort through thousands of codes. This "clinical grouper" makes it easier to quickly understand patterns of diagnoses and procedures to analyze costs, utilization, and outcomes associated with particular illnesses and procedures. We used the single-level CCS which is most useful for ranking diagnoses and procedures and for direct integration into risk adjustment and other software. An excel spreadsheet summarizing CCS diagnosis cross-classifications is available on request.

- **CCS-Services and Procedures** (CCS-SP) software classifies CPT and Healthcare Common Procedure Coding System (HCPCS) codes into clinically meaningful procedure categories.[5] These categories are identical to the ICD-9 procedure CCS with the addition of specific categories unique to the professional service and supply codes in CPT/HCPCS. CCS-SP can be used with any data that include CPT or HCPCS procedure information, in the present instance, the ED/AS data. An excel spreadsheet summarizing all CCS procedure cross-classifications is available on request.

- **Inpatient Quality Indicators** (IQI) are a set of measures used to provide a perspective on quality.3 They include mortality indicators for conditions where mortality has been shown to vary across institutions and for which there is evidence that high mortality may be associated with poorer quality of care. They also include utilization indicators that examine procedures whose use varies significantly across hospitals and for which questions have been raised about overuse, underuse, or misuse. High or low rates for these indicators are likely to represent inappropriate or inefficient delivery of care. For this study we focused on two IQI: Cesarean sections (CSEC) and Vaginal Birth after a previous Cesarean section (VBAC). The software we used to do these calculations is available on request.

Figure 1 summarizes major steps to identify, classify, and categorize events for this study. Source master files were previously prepared, using methods described elsewhere.[6] After pregnancy-related records were identified (CLS), they were categorized (CAT) to prepare for analysis and/or summary.

Figure 1. Classify and Store Data Blocks in Specific Subsets (* = PDD, ASC, ED)



In the PDD, we identified pregnancy-related discharges based on MDC 14 (Pregnancy/Delivery) or DRGs 370-384. Sometimes pregnant women seek care for other reasons. Examples might include injury, asthma, or mental illness. In such events, the principal diagnosis will be something other than pregnancy, but a secondary diagnosis will show the woman as pregnant. Thus we used the CCS to search over the full array of secondary diagnoses to any evidence of the high-level CCS diagnosis category 11 (Pregnancy, birth, or puerperium diagnosis) and the high-level CCS procedure category 13 (Obstetric procedures). In this step we also identified the two PQI events CSEC and VBAC.

The ED/AS datasets lack MDC and DRG variables needed to calculate PQI. Thus, we relied on the CCS to identify pregnancy-related diagnoses and procedures in these settings. For the procedures, we used the CCS formats specifically for CPT codes.

Finally, to compare California trends to the nation, we obtained 2001-2005 data for pregnancy-related diagnoses and procedures from AHRQ's Healthcare Cost and Utilization Project (HCUP) website. The Nationwide Inpatient Sample (NIS) is one in the family of databases and software tools AHRQ developed. HCUP's online query system summarized the annual weighted NIS files by the CCS and exported the results into Excel files.[7] HCUP data were summarized two ways. The first was by principal diagnosis and procedure with outcome statistics. We also obtained frequencies summarized over the array of all available diagnosis and procedure fields. We read the HCUP files into SAS and used them to make various comparative charts and graphs. For national rate trends in CSEC and VBAC deliveries, we used reports published by the National Center for Health Statistics.[8]

Access Indicators

- **Payor.** The variable summarizing expected source of payment, or insurance status, underwent changes between 1983 and 2006, as managed care (MC) took hold. To bridge definition variations, we summarize expected source of payment into the following dichotomous categories: Private (HMO/PHP (managed care, private sector) and employment-related coverage (fee-for-service (FFS), CHAMPUS, workers compensation, and other) and Public (MediCal/Medicare/uninsured). In California, pregnant women who lack insurance are eligible for MediCal and rarely are uninsured once pregnancy is established. MediCare is rarely a payor for a pregnant woman unless she has a pre-existing chronic health condition.

- **Out-of-County.** (OOC) In an environment where hospital availability is decreasing, our concern about out-of-county (OOC) care is three-fold. First, we are concerned about the effect of distance on the ability of the family to access care for their family member, in this instance a pregnant woman. Distance increases the length of time it takes the person to receive care. A time delay may worsen symptoms and negatively impact outcomes. Finally, we are concerned about the burden distance places on the family in the community of residence at a time when bonding is critical. In several previous studies, we have found OOC to have an adverse impact on access and outcomes.[9]

- **Hospital Auspices and Ownership.** (HCAT) In this time of managed care, choice of providers is limited to those offered by the health plan to which the patient subscribes for health insurance coverage. Most people have a limited choice of health plan options, largely governed by what their employer or local county health department offers. In California, subscribers typically "choose" from a group of physicians affiliated with an Independent Practice Association (IPA). The IPA typically is a member of a Provider Healthcare Organization (PHO) composed of itself and a hospital or hospital system. The PHO negotiates with insurers and employers for benefits the PHO members (physicians, hospitals) must provide subscribers. The physician group only can refer patients to hospitals covered by the insurance plan which usually is the PHO hospital member. The only exception is when emergency care is needed. In this instance, the hospital receiving the pregnant woman must treat her. Kaiser Permanente is the originator and largest example of this type of health plan.

  Hospital systems site hospitals and offer (or do not offer) services within hospitals to maximize profits. Hospitals contract with hospitalists to provide inpatient care, including delivery and obstetrics. Hospitalists typically are IPA members who receive higher compensation than other IPA members. Physicians and hospitals split profits through the PHO.

  This is the constricted environment in which pregnant California women "choose" their physicians and the hospitals where their children will be born. Thus, hospital systems available in a given woman's community and the way the hospital system placed its services are a strong factor in access to hospital care.

  For this study, we assigned hospitals based on auspices and system ownership: Non-Profit (Kaiser Permanente (KAI), Sutter (SUT), Religious (REL), and all other non-profits (NFP)), Public (District (DIS), City or County (PUB)), University (UNI), and Investor-owned (INV).[10]

Resource Utilization

- **Length of Stay.** LOS = Discharge date minus admission date. If the patient is admitted and discharged on the same day (LOS = 0), the number is incremented by one.[11] LOS also has a value of one when a patient is admitted on Day 1 and discharged on Day 2. When LOS is longer than 365 days, it is truncated, because reported total charges are based on services during the preceding 12-month period.[12]

- **Total Charges.** Charges are missing non-randomly, because OSHPD exempts some hospitals such as Kaiser and children's hospitals from reporting this. However, charges are reported when non-Kaiser members receive care in Kaiser facilities or Kaiser members receive care in non-Kaiser facilities. In general, reported charges are higher that actual reimbursements, and do not provide a clear picture of the cost of hospital care. However, charge information provides a sense of the relative cost of care, thus allowing comparison between groups of cases. To better estimate the total economic burden of mental health hospitalization, we imputed charges for records lacking them. The method of imputing charges is described elsewhere.[6]

Adverse Outcomes

- **Extended Length of Stay.** (ELOS) Many researchers have investigated ELOS as an adverse outcome indicator.[1 9 13] We did a series of analyses to examine frequencies for stays greater than or equal to 4, 5, or 7 days. Overall, about 87% of discharges ended by 4 days. In the general acute care setting, 85% ended by then. Based on these results, we define a stay of 4 days or longer as extended (ELOS4).

- **Non-routine Disposition**. (NRD) Rather than attempting to measure patient outcomes in terms of specific procedures relative to specific medical conditions, we use the fact of a non-routine discharge disposition (NRD) as a provisional indictor of overall patient discharge quality.[13] This reflects the fact that the woman did not exit the hospital to return home, but transferred to another organized healthcare environment or died. In several studies, we have found NRD to be associated with adverse outcomes.[9 14]

- **Death.** This study offered several opportunities to identify maternal deaths. The first was when a death occurred on a discharge record during a pregnancy event. The second was when a woman was admitted to hospital for a non-pregnancy event and died. The third was deaths identified in the Vital Statistics death file. In the analysis reported here, we use a variable DIED (0 = Alive, 1 = Dead) to indicate that a death was recorded in any one of these circumstances. We use the variable DISPND (0 = Alive, 1 = Dead) when discussing deaths in an OSHPD facility.

- **Patient Safety Indicators** (PSI) are a set of measures that screen for adverse events that patients experience as a result of exposure to the health care system. These events are likely amenable to prevention by changes at the system or provider level.[15] PSI cannot be calculated on the ED/AC datasets because they lack MDC and DRGs. The software we used to do these calculations is available on request. The ED/AS datasets lack MDC and DRG variables needed to calculate PSI. For these settings, we were able to classify adverse events using the small subset of diagnoses and procedures that by definition are adverse events. PSI were calculated as a separate CAT step, using discharges previously identified.

  Table 1 shows the PSI we calculated, with their inclusion and exclusion criteria as defined in the source documents. Note that the list has more than explicitly pregnancy-related PSI. Pregnant women could be admitted for other reasons and have adverse events during those stays. PSI are rare events and most had very small occurrences. Because of variation in criteria, a patient can be included in or excluded from more than one PSI. In addition to the individual PSI, we made a single summary variable flagging the occurrence of any PSI or any discharge with an adverse event by definition.

Table 1. Patient Safety Indicators

Indicator	Definition and Numerator	Denominator
Complications of anesthesia	Discharges with ICD-9-CM diagnosis codes for anesthesia complications] in any secondary diagnosis field per 100 discharges.	All surgical discharges. Exclude patients with codes for poisoning due to anesthetics *E855.1, 968.1-4, 968.7* AND any diagnosis code for active drug dependence, active nondependent abuse of drugs, or self- inflicted injury.
Death in low mortality DRGs Indicator is stratified in 7 subgroups: 1. Adult surgical 2. Adult medical 3. Pediatric surgical 4. Pediatric medical 5. Psychiatric 6. Obstetric 7. Neonatal	All discharges with disposition of "deceased" per 100 population at risk.	All discharges in DRGs with less than 0.5% mortality rate, based on NIS 1997 low mortality DRG. If a DRG is divided into "without/with complications" both DRGs must have mortality rates below 0.5% to qualify for inclusion. Exclude patients with any code for trauma, immunocompromised state, or cancer.
Decubitus ulcer	Discharges with ICD-9-CM code of 707.0 in any secondary diagnosis field per 100 discharges.	All medical and surgical discharges. Include only patients with a length of stay of more than 4 days. Exclude patients in MDC 9 or patients with any diagnosis of hemiplegia, paraplegia, or quadriplegia. Exclude patients admitted from a long term care facility.
Failure to rescue	All discharges with disposition of "deceased" per 100 population at risk.	Discharges with potential complications of care listed in failure to rescue definition (e.g., pneumonia, DVT/PE, sepsis, acute renal failure, shock/cardiac arrest, or GI hemorrhage/acute ulcer). Exclusion criteria specific to each diagnosis. Exclude patients transferred to or from acute care facility or admitted from long-term care facility.
Foreign body left in during procedure	Discharges with ICD-9-CM codes for foreign body left in during procedure] in any secondary diagnosis field per 100 surgical discharges.	All medical and surgical discharges.
Iatrogenic pneumothorax	Discharges with ICD-9-CM code of 512.1 in any secondary diagnosis field per 100 discharges.	All medical and surgical discharges. Exclude patients with any diagnosis of trauma, any code indicating thoracic surgery or lung or pleural biopsy or cardiac surgery.
Infection due to medical care	Discharges with ICD-9-CM code of 999.3 or 996.62 in any secondary diagnosis field per 100 discharges.	All medical and surgical discharges. Exclude patients with any diagnosis code for immunocompromised state or cancer.
Postoperative hip fracture	Discharges with ICD-9-CM code for hip fracture] in any secondary diagnosis field per 100 surgical discharges.	All surgical discharges. Exclude patients who have musculoskeletal and connective tissue diseases (MDC 8); with principal diagnosis codes for seizure, syncope, stroke, coma, cardiac arrest, poisoning, trauma, delirium and other psychoses, or anoxic brain injury; any diagnosis of metastatic cancer, lymphoid malignancy or bone malignancy, self-inflicted injury, 17 years of age and younger.
Transfusion reaction	Discharges with ICD-9-CM codes for transfusion reaction] in any secondary diagnosis field per 100 discharges.	All medical and surgical discharges.
Obstetric trauma - vaginal with instrument	Discharges with ICD-9-CM codes for obstetric trauma] in any diagnosis or procedure field per 100 instrument assisted vaginal deliveries.	All vaginal delivery discharges with any procedure code for instrument assisted delivery.
Obstetric trauma - vaginal without instrument	Discharges with ICD-9-CM codes for obstetric trauma] in any diagnosis or procedure field per 100 instrument assisted vaginal deliveries.	All vaginal delivery discharges patients. Exclude instrument assisted delivery.
Obstetric trauma - cesarean section	Discharges with ICD-9-CM codes for obstetric trauma] in any diagnosis or procedure field per 100 cesarean deliveries.	All cesarean delivery discharges.

Prepare Data to Identify "Women"

Unique Identifier

The datasets we used are confidential. This means they contain protected health information and require special research protocols. FHOP's protocol permits us to use the original SSN rather than OSHPD's Record Linkage Number. We encrypt the SSN with our own algorithm to facilitate soft linkages using a combination of probabilistic and deterministic methods. The resulting variable is called SSNC.

OSHPD's California Hospital Outcomes Project did not do soft linkages in its first study of delivery outcomes.[1] [16] We agree with the reasoning of that study. The sheer volume of deliveries, combined with the large numbers of women who lack SSN and the difficulties sequencing episodes of care for this population led to the decision we followed here. Many records lacked SSNC (21.6%) compared with OSHPD's earlier study, 20.1%.[1]

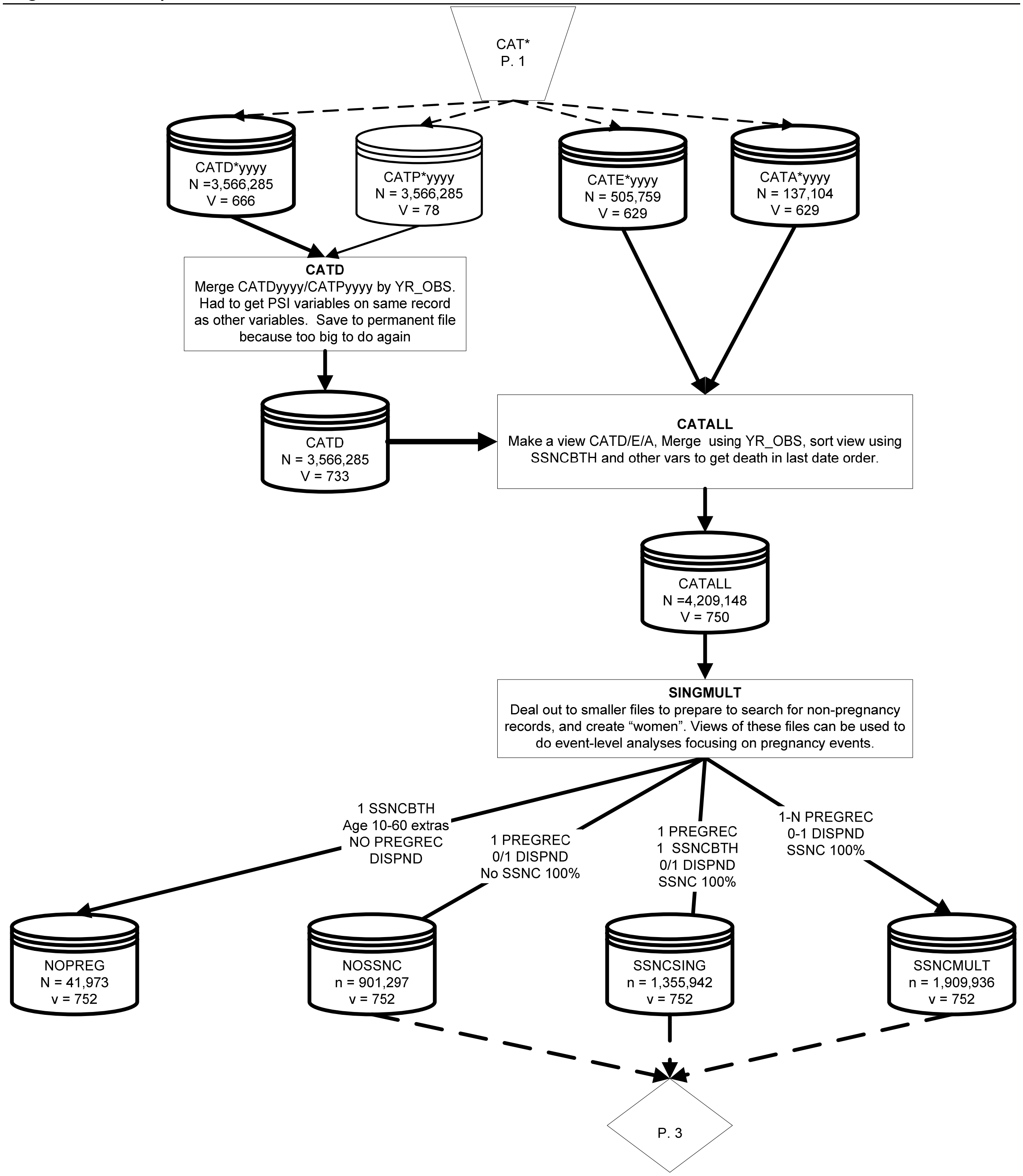
We concatenated the encrypted Social Security Number (SSNC), sex (in this case female) and birth date to create a unique identifier (SSNCBTH). If 2 records had the same SSNC but different birth dates, they were considered as two women. We did not discard the second birth date record, but thought in linking to the death certificates where there was a death, we would find one or the other. When SSNC was missing, SSNCBTH consisted of birthdate, sex, ZIP-code of residence, and race/ethnicity.

Merge Data and Output Based on Unique ID Characteristics

Figure 2 shows the data flow during this preparatory step. thicker lines indicate "fat" files with many variables. We first put the various PDD files together (CATD), merged that result with the ED/AS data (CATALL), and dealt them out to four very large files (SINGMULT):

- NOPREG - Death events that did not link to a pregnant woman (N = 41,973).
- SSNCSING - SSNC with one event (N = 1,355,942).
- SSNCMULT - SSNC with more than one pregnancy-related event (N = 1,909,936).
- NOSSNC - Events lacking SSNC (N = 901,297). SSNC were missing non-randomly. Women lacking SSNC were more likely to be Hispanic (84%), treated inpatient (84%), in public (43%) and investor (32%) hospitals. Women age 15 to 24 were more likely not to have a SSNC (26%) compared with women 35 and older (13%). SSNC was missing for 19% of Bay Area records and 22% of records from elsewhere in California. Linkage was attempted only for death events (N = 141).

Figure 2. Prepare data to find "women"



Calculate the Pregnancy Window

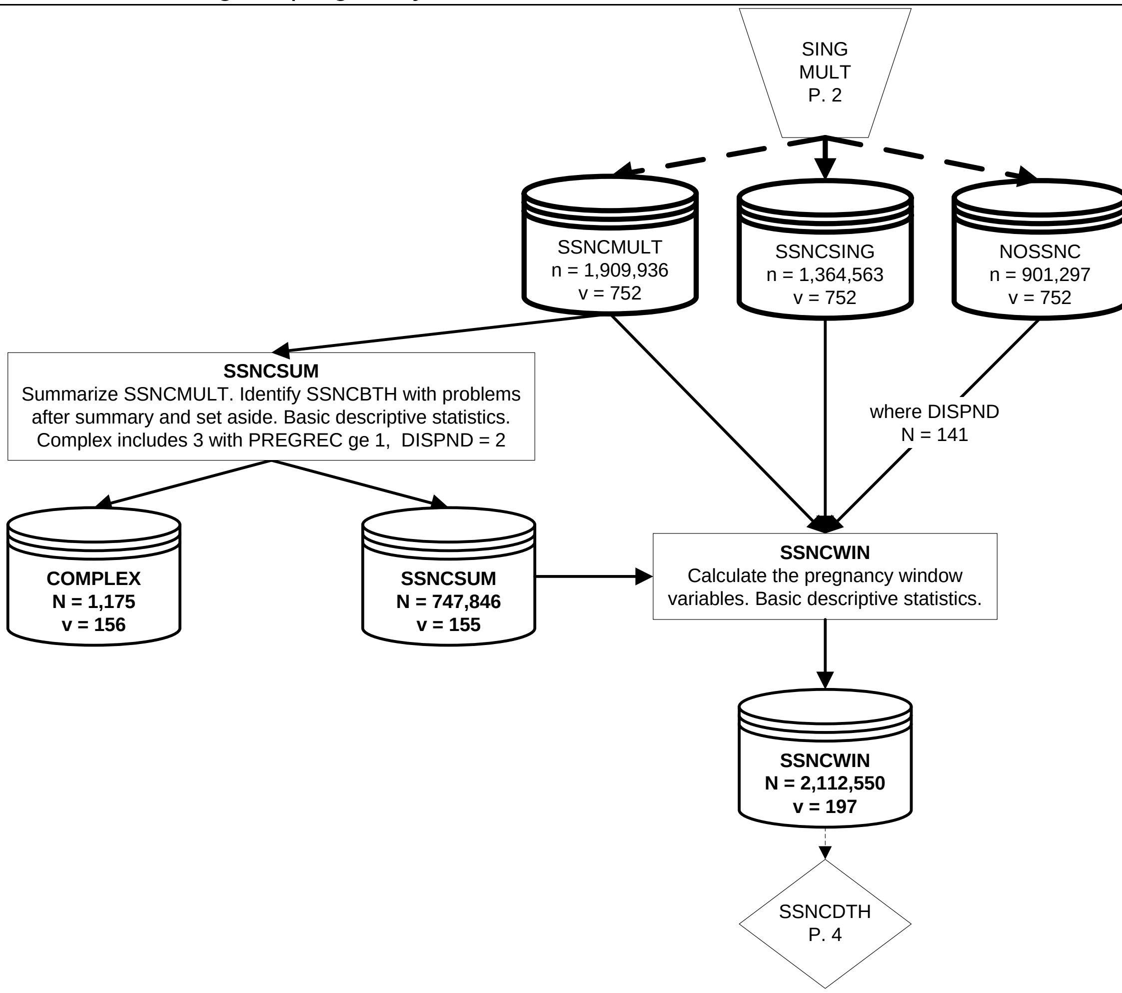
We defined the pregnancy window as the admission date for a woman's first pregnancy-related event through the discharge date for her last pregnancy-related event. Over the six year study period, a woman could have been pregnant more than once. She could have been admitted 1 to N times for each pregnancy. For discharges lacking SSNC (Fig. 2, NOSSNC), we had no way to know how many women were represented and how many records represented the same woman. We made made no effort to link these records. Except for those with deaths, records in NOSSNC were set aside. Women with SSNC who had only one discharge (Fig 2, SSNCSING) by definition had one pregnancy-related discharge (typically a delivery). We knew how many women were in these single-event groups.

We did not know how many women were in the set with multiple events (Fig. 2, SSNCMULT). Thus our next step involved summarizing that dataset. We also calculated reliability statistics. For example, did the woman have more than one race/ethnic group recorded over the set of records. After reviewing reliability statistics, we set aside complex cases (e.g., 3 or more races for the same SSNC or 72 pregnancy records). In the final dataset, 99.6% had 1 race over the 2 to n records. When this was completed, we merged the summary file with the single SSNCBTH file and added death records from the file lacking SSNC.

To describe outcomes and set the stage to calculate reliability for death linkages, we also obtained the DRG, E-Code, summary PSI indicator, ZIP and county of residence, and county of discharge for the first and last pregnancy event, and if the woman died, the death event. Note that the death can have occurred during a pregnancy event, delivery event, or post-delivery at any time from last pregnancy or delivery until death. Also note that all women who were pregnant did not die in one of the studied settings. Thus we did not have our best count of maternal deaths until after merging with the death files.

Figure 3 summarizes work to calculate the pregnancy window. From SSNCMULT, we discarded 1,176 women with data complexities (COMPLEX) and retained 747,846 women who had more than one pregnancy event (SSNCSUM). We merged this file by SSNCBTH with SSNCSING and NOSSNC (the 141 cases with deaths) to create the SSNCWIN (N = 2,112,550) file of women for whom we could try to link to the death file.

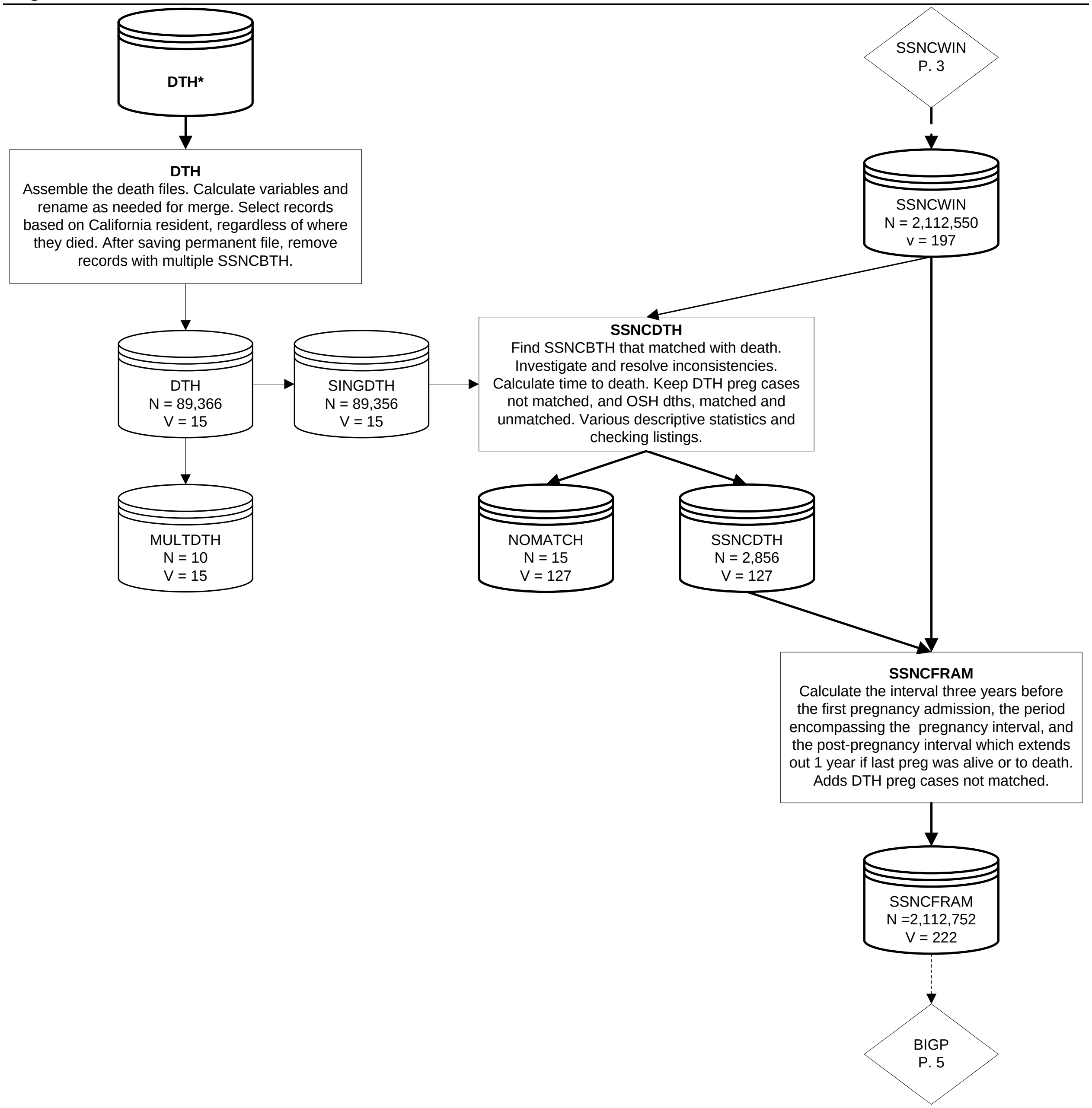
Figure 3. Calculating the pregnancy window



Link Women to Death File

Before linking discharges to deaths, we had to prepare the death file. This involved identifying all women age 10 to 60, and calculating or renaming variables in these files that were needed for the linkage. In addition to age and sex, we selected records based on California residence regardless of where the woman died. Hoping to link the few women who died and lacked SSNC, we kept all death records. We removed SSNCBTH with more than one record (n = 10, or 5 sets), resulting in a file of 89,356 deaths between 2001 and 2005. Figure 4 summarizes the steps to complete activities needed to link data to the death records and prepare the data to find non-pregnancy records. The file SINGDTH is ready for linkage.

Figure 4. Link Women to Death File



To do the linkage, we used a macro that makes a unique match given certain criteria. If multiple records meet the criteria, no selection is made (Fig 4, SSNCDTH). The macro matches one-to-one, and not one-to-many. Among other applications, the macro was used for FHOP's adolescent injury study[17] and more recently for OSHPD's Intensive Care Outcomes study.[18]

Because the 2006 death file did not arrive in time, we were unable to search for out-of-hospital deaths for women discharged in 2006, or for one-year mortality for women discharged in 2005. After linkage, two residuals remained: women who hospitals reported as having died, and women from the death file whose underlying cause of death was pregnancy-related and who reportedly died in a hospital setting. We merged these two files and manually reviewed them to evaluate if matches were possible. None were found.

Table 2 summarizes linkage results by linkage algorithm. We had three sets of "women": those OSHPD recorded as in-patient deaths that were found in the death file (OSHPD = Yes, DEATH = Yes, N = 2,117), those discharged as dead but not found (OSHPD = Yes, DEATH = No, N = 537), and those the death file said died in a hospital setting of pregnancy-related causes that were not found in the OSHPD file (OSHPD = No, DEATH = Yes N = 202). Most deaths in the second circumstance were primarily 2006 deliveries when we did not have 2006 deaths. Most of the latter deaths were before 2005, when the ED/ASC datasets were not available.

Table 2. Results of linking OSHPD and Death files

Data Source		Have	Linkage Algorithm				Linked		Reliability
OSHPD	DEATH	SSNC	SSNC	BirthDt	DeathDt	ZIP	N	% Total	Score
Total							2,856	100.00	
Yes	Yes	Yes	Exact	Exact			1,934	67.72	5.8
			Exact		Exact		22	0.77	5.5
			Exact	Soft			54	1.89	4.1
			Soft	Exact		Exact	35	1.23	5.3
				Exact	Soft	Exact	5	0.18	6.0
			Soft		Exact	Exact	1	0.04	6.0
				Soft	Exact	Exact	1	0.04	6.0
		No		Exact		Exact	44	1.54	7.0
				Exact	Exact		15	0.53	4.5
				Exact		Exact	1	0.04	5.0
					Exact	Exact	5	0.18	5.8
Yes	No	Yes					461	16.14	
		No					76	2.66	
No	Yes	Yes					152	5.32	
		No					50	1.75	

This table identifies elements of the linkage algorithm. "Soft" means we allowed for possible errors on one side or the other (e.g., off by one element of birth or death date). We did not use soft matches for the 141 OSHPD death cases lacking SSNC. For those cases, SSNCBTH by definition lacked SSNC but included race/ethnicity, birth date, ZIP of residence, and sex.

We calculated various agreement statistics to check linkage reliability. To be kept as a linkage, we required a minimum reliability score of 2. We set aside 15 linkages did not match on at least 2 of our linkage variables (Fig 4, NOMATCH). None of these had been shown as deaths in the OSHPD file. The last column in Table 2 shows average linkage agreement by method.

We used linked cases (N = 2,117) to calculate overall reliability. Because the analysis has a geographic focus (Bay Area or elsewhere in California), we tested if women we linked had a geographic relationship to the area where they delivered and died. Thinking that time to death (at delivery through 1 year post-delivery versus death at a later date) would affect geographic statistics, we calculated reliability tests by time to death. Agreement above 75% may be taken to represent excellent agreement beyond chance.[19] Table 3 summarizes these results.

Table 3. Linkage Agreement (%) by Time to Death

		Time to Death	
Agreement Percent	Overall	1 year	> 1 Year
Number of women	2,117	1,070	1,047
SSNC+Sex+Birthdate	93.4	93.4	93.5
SSNC only or ZIP5+Sex+Birthdate	97.2	96.6	97.8
Race/Ethnicity	86.6	86.0	87.2
County where woman died	96.6	97.9	95.2
County where woman lived	95.0	96.3	93.6
ZIP3 where woman lived	89.7	92.1	87.2
ZIP5 where woman lived	72.8	77.8	67.7
In-Hospital death	44.1	55.9	32.1
Overall average agreement	5.8	5.9	5.5

We found no significant time-based differences for linkages based on SSNC, its related combinations, or race/ethnicity.

Overall geographic reliability was high but differed based on time to death (P <0.001). Agreement was lowest for in-hospital death.

Agreement scores ranging from 2 to 7, and 95% agreed on 4 or more elements. The average agreement score was higher for women who died within 1 year of last pregnancy discharge than for those who died later (t = 8.31, P = <0.001). Thus we are fairly confident in our linkage, even after giving women the opportunity to move.

We added to our window file those women in the death file we had not found earlier, calculated time to death following the last pregnancy-related discharge, and calculated a frame around the window within which to find non-pregnancy records or records for the death file women we had not previously identified (Fig 4, SSNCFRAM, N = 2,112,752). The frame was defined as three years before the first pregnancy record through one year post-discharge or death whichever came last. Note that the number of cases is incremented to include the additional deaths not previously found and for which we hoped to find discharge data.

Find Non-Pregnancy Records

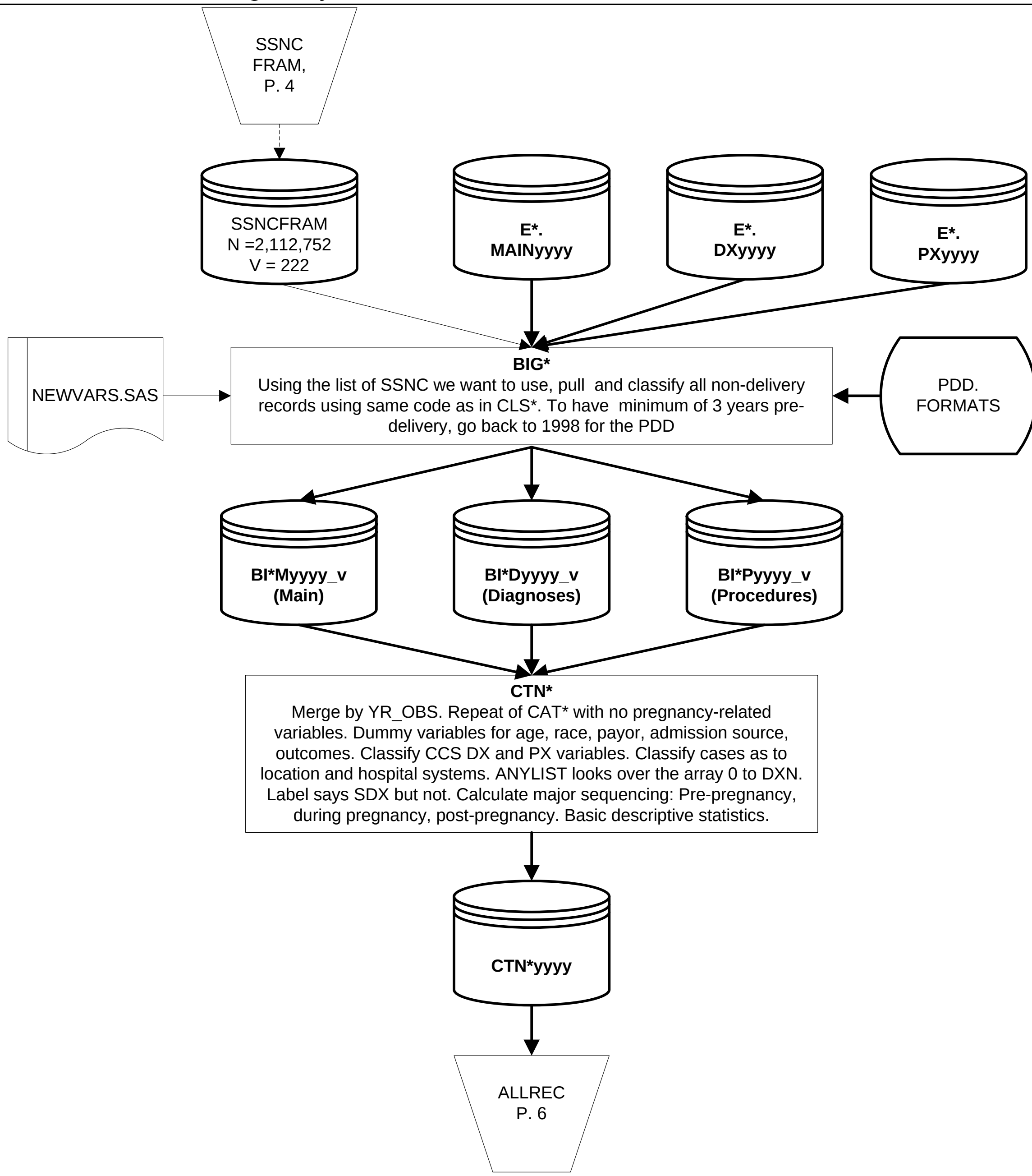
We were interested in finding non-pregnancy records because they may provide a fuller description of a woman's life course with respect to health conditions that may underlie an adverse delivery outcome or death and that might not be recorded on delivery records. These could include conditions of uncertain clinical significance for pregnancy, antepartum conditions that resolve before delivery,[16] or post-partum conditions that may not be reflected in the Underlying Cause of Death codes. Examples include mental illness or substance abuse

(MI/SA), admissions for ambulatory-care-sensitive (ACS) diagnoses such as epilepsy or asthma that could indicate having problems accessing outpatient care, or injuries. Whether these occurred before, during, or after the pregnancy window they may have had lingering impacts on a woman's health during a time when she was the mother of an infant or young child.

These circumstances arise partly because ICD-9-CM coding rules are subject to interpretation and may not mandate coding. Hospitals are required to report only "conditions that affect patient care in terms of requiring: clinical evaluation; or therapeutic treatment; or diagnostic procedures; or extended length of hospital stay; or increased nursing care and/or monitoring."[20] Conditions may be coded more faithfully when they represent an actual indication for delivery because "diagnoses that relate to an earlier episode which have no bearing on the current hospital stay are to be excluded."[16] Hence, conditions such as MI/SA, ACS, or injury must be reported only if active at delivery ***and*** the condition affects obstetric management. A history of problems accessing healthcare or MI/SA will not be fully identified by focusing only on pregnancy records.

Figure 5 shows the series of programs to find non-pregnancy admissions. These programs essentially repeated the sequence in Figure 1 to find, classify, and categorize delivery-related records. The difference in this step was that we had the SSNCBTH we were seeking, knew the frame dates, and could exclude previous records found. We did not seek additional records of women lacking SSNC. The abbreviation BI* stands for what we call our "big pull". CTN* is an abbreviation indicating the step where we categorize new records from each source and year.

Figure 5. Find Non-Pregnancy Records



Summarize Data to a "Woman"

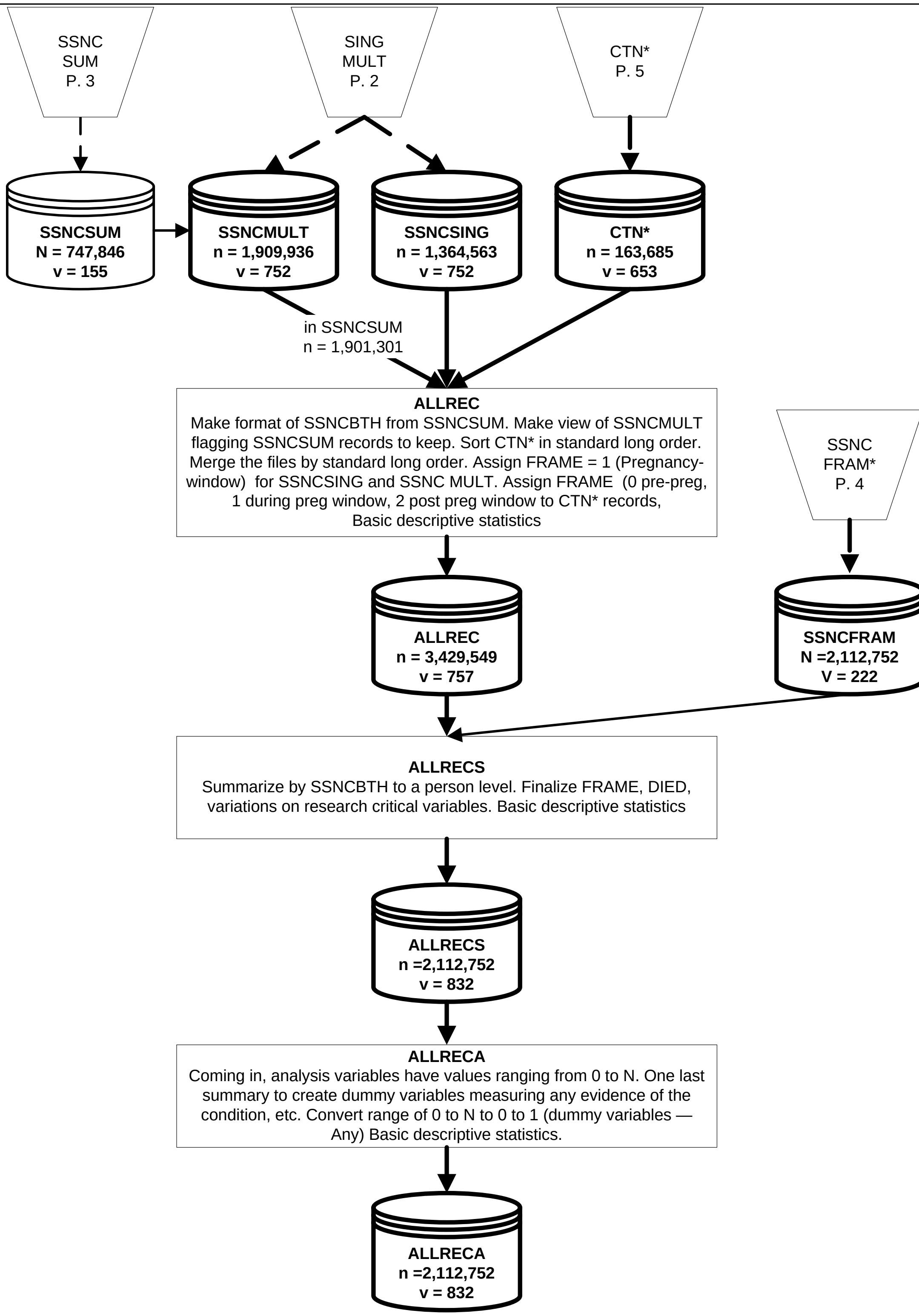
We did not analyze the data based on an episode of care (EOC), the period initiated by patient presentation with a diagnosis of a clinical condition and concluded when the condition is resolved.[21] For example, the California Hospital Outcomes Project defined an EOC for a heart attack as the period from first admission to an acute care hospital for a heart attack to final discharge within 30 days, including transfers or readmissions to another hospital.[22] OSHPD's first delivery outcomes study linked delivery records with both postpartum (within 6 weeks after delivery) and antepartum hospitalizations (within 273 days before delivery).[1] These studies deleted all records between EOC that were not relevant to the condition studied while we explicitly wanted to retain them.

Linking EOC is an expensive and time-consuming process. We did not have the time or resources to do this. Instead, we summarized linked records by SSNCBTH, or "woman", to gain a high-level view of her health over that part of her reproductive history included in the 6-year pregnancy window 2001-2006 and the pregnancy frame 3 years before her first pregnancy admission and 1 year after her last pregnancy discharge or death.

We began by pulling all records found into a large file by discharge event (ALLREC). This file could be summarized into EOC because of the way we merged the varying data sources. Because ALLREC was so large (N = 3,429,549), we had to break it into five parts (not shown) for summary. The file ALLRECS (N = 2,112,752) is the result. In this file, a given indicator can have a value ranging from 0 to N, depending on the number of records where the condition was flagged. In the file ALLRECA, the indicator data was further summarized to identify the occurrence of any dummy variable (0 = condition not found, 1 = condition found 1 or more times). Figure 6 summarizes these steps.

We assigned a single race/ethnicity and age category for analyses. If the woman had the same race/ethnicity in both files, we used that. If they were discrepant, we gave first priority to American Indian/Native American on either side, then Hispanic on either side, then Black on either side, then Asian. We assigned age categories as follows for each admission: 15 to 24, 25 to 34, 35 to 45. Small numbers of women below or over those ages were assigned respectively to be 15 to 24 and 35 to 45. With a six-year pregnancy window, a woman could have been age 21 at her first pregnancy and 28 at her last. Where ages spanned categories over the period, we assigned the last age category at her last pregnancy. In the example, the final age category would be 25 to 34.

Figure 6. Summarize Data to a "Woman"



Summary

We have summarized the complex steps to create different types of analysis files: discharge-level files for all pregnancy-related conditions, discharge-level files for all records for women with a SSNC, and summary files for "women" with SSNC. The pregnancy window was defined as the period between the admission date for the first pregnancy-related event and the discharge date for the last pregnancy-related event, between 2001 and 2006. The pregnancy frame was defined as three years before the first pregnancy-related event and one year after the last pregnancy-related event or death. Deaths are under-counted because we did not have the 2006 or 2007 death files to calculate one-year mortality (pregnancy-associated death). All work was done in SAS, and programs are available upon request.

A separate standard sequence of FHOP programs summarized delivery records for the pregnancy indicators cesarean delivery, primary cesarean delivery, and any adverse event. A description of these indicators, how they were calculated and their use is available elsewhere.[23]

ENDNOTES

[1] Romano PS, Luft HS, Remy LL, Rainwater J, Rennie DJ, Zhou H, et al. Report of the California Hospital Outcomes Project: maternal outcomes following delivery, risk-adjustment methodology and preliminary findings. Sacramento (CA): Office of Statewide Health Planning and Development; 1996.

[2] Hornbrook M, Whitlock EP, Berg CJ, Callaghan WM, Bachman DJ, Gold R, Bruce FC, Dietz PM, Williams SB (2007) Development of an algorithm to identify pregnancy episodes in an integrated health care delivery system. Health Services Research 42:2 (April 2007).

[3] Guide to Inpatient Quality Indicators: Quality of Care in Hospitals – Volume, Mortality, and Utilization. Department of Health and Human Services Agency for Healthcare Research and Quality June 2002 Version 3.1 (March 12, 2007) Last accessed 07-Dec-2007 at: http://www.qualityindicators.ahrq.gov

[4] Elixhauser A, Steiner C, Palmer L. Clinical Classifications Software (CCS), 2006. U.S. Agency for Healthcare Research and Quality. Available: http://www.ahrq.gov/data/hcup/ccs.htm#download.

[5] Clinical Classifications Software for Services and Procedures. Last accessed: 22Jun2007 at: http://www.hcup-us.ahrq.gov/toolssoftware/ccs_svcsproc/ccssvcproc.jsp#table1

[6] Remy LL, Clay T, Oliva G. (June 2004) Methods to Prepare Hospital Discharge Data. Family Health Outcomes Project, University of California, San Francisco. Available at: http://www.ucsf.edu/fhop/_docs/pdf/pubs/METH_PDDS.pdf

[7] Nationwide Inpatient Sample. Last accessed 03-Jan-2008 at: http://www.hcup-us.ahrq.gov/nisoverview.jsp

[8] 2005 rates: Hamilton BE, Martin JA, Ventura SJ. (2007) Births: Preliminary data for 2005 National vital statistics reports; vol 55 no 11 Hyattsville, MD: National Center for Health Statistics. 2001-2004 rates: Martin JA, Hamilton BE, Sutton PD, et al. (2006) Births: Final data for 2004 National vital statistics reports; vol 55 no 1 Hyattsville, MD: National Center for Health Statistics.

[9] Remy LL, Oliva G (In review) Hospital Capacity to Treat Mental Illness.

[10] Patient Discharge Pivot Profile (2005). Last accessed 24-Mar-2008 at: http://www.oshpd.ca.gov/HID/Products/PatDischargeData/PivotTables/PatDischarges/default.asp

[11] Patient Discharge Data File Documentation January-December 2005. (July 2006) State of California, Office of Statewide Health Planning and Development, P. 15.

[12] Patient Discharge Data File Documentation January-December 2005. (July 2006) State of California, Office of Statewide Health Planning and Development, P. 26.

[13] DeMoro D. (Sep, 1999) California Health Care: Sicker patients, fewer RNs, fewer staffed beds. Orinda, CA: Institute for Health and Socio-Economic policy. Secondary analysis of data from the Dartmouth Health Care Atlas, 1999. Last accessed 19 Mar 2004 at: http://www.calnurse.org/cna/pdf/StaffingRatios6.pdf

[14] Oliva G, Remy LL, Clay T. (2004). The Impact of Changing Public Policy on California's Hospital Infrastructure and Children's Hospital Outcomes - 1983-2000. University of California San Francisco, Family Health Outcomes Project. See: http://www.ucsf.edu/fhop/_htm/publications/index.htm.

[15] Patient Safety Indicators Fact Sheet. (Feb 2006). A detailed Guide to Patient Safety Indicators, software, and software documentation are available on the AHRQ Quality Indicators web site: http://qualityindicators.ahrq.gov/psi_download.htm.

[16] Yasmeen S, Romano PS, Schembri ME, Keyzer JM Gilberg WM. (2006) Accuracy of obstetric diagnoses and procedures in hospital discharge data. American Journal of Obstetrics and Gynecology. 194, 992–1001.

[17] Remy L, Clay T, Oliva G. (Aug 2000). The California Child and Youth Injury Hot Spot Project Report for the Period 1995 to 1997, Volume Three, Technical Guide. Sacramento, CA: California Department of Health Services. See: http://www.ucsf.edu/fhop/_htm/publications/index.htm.

[18] Dudley RA, Kuzniewicz M, Dean M, Lane R, Rennie D, Bacchetti P, Clay T, Crane S, Luft H. (May 2007) Final Report, California Intensive Care Outcomes (CALICO) Project. Office of Statewide Health Planning and

Development. Last accessed 06-Mar-2008 at: http://www.oshpd.ca.gov/HID/Products/PatDischargeData/ICUDataCALICO/index.html

[19] Fleiss JL. (1981) Statistical Methods for Rates and Proportions, Second Edition. New York: John Wiley & Sons.

[20] Coding Clinic for ICD-9-CM 1990;7:13.

[21] Schulman KA, Yabroff KR, Kong J, Gold KF, et al. (1999) A claims data approach to defining an episode of care. Health Services Research, June 1999, 34(2): 603:621.

[22] Romano PS, Remy LL, Luft HS. (1993). Report of the California outcomes project: Volume 2, Technical Appendix. Sacramento, California: Office of Statewide Health Planning and Development.

[23] Remy L, Clay T, Oliva G (2008) Maternal Quality and Safety Indicators Databook 2001-2006. Available at: http://www.ucsf.edu/fhop\publications.html.