**Research Data Extraction Template**

**IRB Number:**

**PI:** Laura Wherry

**Purpose:** This project addresses the marginal value of medical care for pregnant women and their infants by examining outcomes related to interventions delivered during the prenatal, delivery, or neonatal periods. It will use diagnosis cutoffs for gestational diabetes to examine the short- and long-term consequences of diagnosis for health care utilization and health outcomes for the mother and the infant. It will also examine the effect of a c-section delivery on health and health care utilization for the mother and child following delivery.

**Data Elements:**

*Patient data extracted from UCLA’s clinical record systems will be represented in multiple tables having many-to-one relationships among the observations. These tables will be linked by an encounter ID, patient StudyID, or both when needed.*

**Patient or Encounter Inclusion/Exclusion Criteria**

*Describe the criteria that should be used to include patients (or patient encounters) in the data pull. Generally, criteria should be stated in terms of the variables described in the tables below. If possible, include a time range, e.g. from 7/1/2013 to 6/30/2014 (or ongoing, if needed).*

Please list your selection criteria below:

**Selection Criteria**

2 **de-identified** data extractions: 1 for the mother population and 1 for the child population:

1. All available records for any female with a pregnancy-related diagnosis code (ICD-9: 630-679, V22-V23; ICD-10: Z34, Z3A, Z37, O categories) at any time during the period 1/2006 to the present
2. All available records for any children who were the result of the pregnancies of the women identified under (1) at any time during the period 1/2006 to the present

Please provide years and relative dates (days from 1st encounter)

**Result tables** – Some variables without a checkbox are extracted automatically and are required to be pulled. Every table indexed by Unique Patient Study ID.

**Patient Demographics** (one row per patient)

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | | Include Variable | Indicate or describe values to include |
| **Study ID1** | |  |  |
| Age | |  |  |
| Sex | |  |  |
| Race | |  |  |
| Ethnicity | |  |  |
| Vital Status2 | |  | *Both* not known deceased and known deceased |
| Neighborhood ADI category | |  |  |
| Current PCP ID | |  |  |
| Last Visit Encounter Year | |  |  |
| Other (write in) 3 |  |  | |
|  |  | Health insurance plan  Indicator for whether PCP at UCLA  First visit encounter year | |

1 Study ID is a study-specific one-way hash of the unique research identifier assigned to every UCLA patient. The Study ID is used implicitly in each table below to represent the patient identity.

2 Vital status is not known deceased or deceased status in EHR. Note only in-hospital death is recorded, for the most part.

3 Other possible variables include: Marital status, Primary Language, Religion, etc.

**Patient Identifiers** (one row per patient) *– for ease of post-study de-identification*

|  |  |  |
| --- | --- | --- |
| Variable | Include Variable | Indicate or describe values to include |
| Study ID |  |  |
| MRN |  |  |
| DOB1 |  |  |
| Name |  |  |
| Address |  |  |
| 5-digit Zip2 |  |  |
| Phone |  |  |
| Email3 |  |  |
| Other identifiers (write in) |  |  |

Footnotes:

1 Requires data use agreement (DUA) for limited dataset

2 For de-identified dataset, we will provide only the first 3-digits of the zip code for geographic areas with populations > 20K.

3 Email address available only for patients who use MyChart (currently about 40% overall at UCLA). Email is not a secure mode of communication and contacting patients by email may not be permitted for some studies.

**Encounters** (including hospitalizations, outpatient visits, and other encounter types)

Data is readily available from January 1, 2006. Data prior to this date may be requested but will take longer to extract.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variable | Include Variable | | | Indicate or describe values to include | | |
| Study ID/Encounter ID (i.e., Study CSN) |  | | | |  | |
| Epic Encounter Type1 | |  |  | | | |
| Encounter Date (or Days since 1st Encounter) | |  | Days since 1st encounter; De-identify dates (Provide year and relative dates e.g. days since 1st encounter); Please include all data from 1/2006 to Present | | | |
| Encounter Age | |  |  | | | |
| Admit Date | |  |  | | | |
| Discharge Date | |  |  | | | |
| Hospital Discharge Disposition | |  |  | | | |
| ED Disposition | |  |  | | | |
| PCORnet Visit Type2 | |  |  | | | |
| Visit Provider ID3 | |  |  | | | |
| Epic Department Name/ID4 | |  |  | | | |
| Department Specialty | |  |  | | | |
| Location (RR, SM etc.) | |  |  | | |
| Other (write in) | |  |  | | |

Footnotes:

1 Values as recorded in Epic: Office Visit, Hospitalization, Surgery, etc.

2 Standardized visit types from the PCORnet: ED, ED to IP, IP, Ambulatory visit, Non-Acute institutional stay, Other ambulatory visit, No information, Unknown, Other

3 Sensitive data, potentially available in LDS with plan for protection and IRB approval or the values can be coded;

4 In Epic, physical location of care

**Provider**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Include Variable | | | Indicate or describe values to include |
| Provider ID1 | |  | | Encoded so that data are de-identified |
| Provider Name | |  | |  |
| Provider Gender | |  | |  |
| Provider Type | |  |  | |
| Primary Specialty | |  | |  |
| UCLA Employee Flag | |  | |  |
| Other (write in)2 | |  | |  |

Footnotes:

1 This could link to either a Visit Provider ID (in the Encounter table) or to a Current PCP ID (in the Patient Demographics table). May need to be encoded if providers need to be de-identified

2 Other possible variables include: NPI, DEA, Year of Birth, etc. is available if needed

**Hospital Unit Transfers** (Within-hospital patient movement events: Admit, Discharge, Transfer time and location data to the bed and department level if needed)

Data is readily available from March 2, 2013. Data prior to this date is unlikely to be available.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Include Variable | | | Indicate or describe values to include | |
| Study ID/Encounter ID | | |  | |  |
| ADT Event Type1 | |  | |  | |
| ADT Event Date/Time2 | |  | |  | |
| Epic Department Name/ID | |  | |  | |
| Department Specialty | |  | |  | |
| Location (RR, SM etc.) | |  | |  | |
| Other (write in)3 | |  | |  | |

Footnotes: **If any data is requested in this table, an Encounter table containing hospitalizations shall be provided.**

1 Admit, Transfer, Discharge, All

2 Sensitive data, potentially available in limited dataset with plan for protection and IRB approval

3 Other possible variables include: Bed Number, Patient Service on Event, etc., if needed

**Diagnoses** (Encounter billing diagnoses are extracted unless otherwise specified in the Diagnosis Source – problem list diagnoses can also be included; see footnote 1). Typically, we provide ICD-9 prior to 10/01/2015 and ICD-10 after that date.

Some data is available starting in 1991.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Include Variable | | Indicate or describe values to include | |
| Study ID/Encounter ID | |  | |  |
| Diagnosis Source1 |  | |  | |
| Diagnosis Date/Time |  | | De-identify dates (Provide year and relative dates e.g. days since 1st encounter); Please include all data from 1/2006 to Present | |
| ICD Type2 |  | |  | |
| ICD Code |  | |  | |
| ICD Description |  | |  | |
| Primary Diagnosis Flag 3 |  | |  | |
| Admission Diagnosis Flag 4 |  | |  | |
| Present on Admission 5 |  | |  | |
| Hospital Final Diagnosis 6 |  | |  | |
| Other (write in) 7 |  | |  | |

Footnotes:

1 Possible values: Ambulatory Encounter Diagnoses, Hospital Final (Discharge) Diagnoses, Hospital Admission Diagnoses. Hospital or Professional Billing Diagnoses are available on special request

2 Possible values are 9 or 10 which signifies ICD-9 or ICD-10 respectively. ICD-10 data not available prior to CareConnect go-live date (03/01/2013)

3 Values include P (primary) or S (secondary), or Null.

4 Values include A (admit) or Null.

5 Value will be Y is diagnosis is present on admission.

6 Value will be 1 if diagnosis is a final hospital discharge diagnosis.

7 Other possible variables include: SNOMED-CT code and description etc.

**Procedures Completed** (per billing data, with CPT, ICD-9 procedure codes, or other codes)

Some data is available starting in 2006.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Include Variable | | Indicate or describe values to include | |
| Study ID/Encounter ID | |  | |  |
| Procedure Date |  | | De-identify dates (Provide year and relative dates e.g. days since 1st encounter); Please include all data from 1/2006 to Present | |
| Procedure Type1 |  | | CPT, ICD-9, or ICD-10 | |
| Procedure Code2 |  | |  | |
| Procedure Description |  | |  | |

1ICD-9, ICD-10, CPT-Professional, CPT-Hospital and CPT-preCC

2ICD-9 procedure codes end and ICD-10 procedure codes start 10/01/2015.

**Problem List**

Data is only available from March 2, 2013. Data prior to this date is not available. Typically, we provide ICD-9 prior to 10/01/2015 and ICD-10 after that date.

|  |  |  |
| --- | --- | --- |
| Variable | Include Variable | Indicate or describe values to include |
| Problem List ID |  |  |
| ICD Type1 |  |  |
| ICD Code |  |  |
| ICD Description |  |  |
| Problem Description |  |  |
| Noted Date |  |  |
| Date of Entry |  | De-identify dates (Provide year and relative dates e.g. days since 1st encounter); Please include all data from 3/2013 to Present |
| Deleted Date |  |  |
| Resolved Date |  |  |
| Problem Status |  |  |
| Problem Comment |  |  |
| Priority |  |  |

Footnotes:

1 Possible values are 9 or 10 which signifies ICD-9 or ICD-10 respectively. ICD-10 data not available prior to CareConnect go-live date (03/01/2013)

**Appointments** (Completed appointments are in the Encounters table. This table is only useful if cancelled, no-show, future appointments, etc. are needed)

Data is only available from March 2, 2013. Data prior to this date is not available.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Include Variable | | | Indicate or describe values to include |
| Study ID/Encounter ID | |  |  | |
| Appointment Date and Time | |  | |  |
| Appointment Made Date and Time | |  | |  |
| Department ID | |  | |  |
| Department Name | |  | |  |
| Department Specialty Name | |  | |  |
| Center Name | |  | |  |
| Location Name | |  | |  |
| Provider Name with ID | |  | |  |
| Referring Provider Name with ID | |  | |  |
| Procedure Name | |  | |  |
| Appointment Confirmation Status | |  | |  |
| Appointment Status1 | |  | |  |

Footnotes:

1 Values are: Scheduled, Completed, Canceled, No show, Left without seen, Arrived.

**Vital Signs and Other Flowsheet Data**

Data is readily available from March 2, 2013. Data prior to this date is unlikely to be available.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | | Include Variable | | Indicate or describe values to include |
| Study ID/Encounter ID |  | |  | |
| Vital Sign Type | |  | |  |
| BMI | |  | |  |
| Height | |  | |  |
| Weight | |  | |  |
| Blood Pressure | |  | |  |
| Pulse | |  | |  |
| Temperature | |  | |  |
| Respiratory Rate | |  | |  |
| O2 Saturation (SpO2) | |  | |  |
| Vital Sign Value | |  | |  |
| Vital Sign Taken Time | |  | | Specify Time Range: De-identify dates (Provide year and relative dates e.g. days since 1st encounter); Please include all data from 3/2013 to Present |
| Other Flowsheet Data(write in)1 | |  | |  |

Footnotes: **Counts of each flowsheet measure can be pulled first for review and to allow for more detailed selection.**

1 Other possible variables include: Pain Score, Ventilator Settings, Inputs and Outputs, GCS Score, and Assessments, ICU monitor data etc.

**Laboratory Test Results**

Data is readily available from 2006. Data prior to this date may be requested but will take longer to extract.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Include Variable | | Indicate or describe values to include | |
| Study ID/Encounter ID | |  | |  |
| Component (Result) ID |  | |  | |
| Component (Result) Name |  | |  | |
| Specimen Taken Time |  | |  | |
| Procedure (ordered) ID 1 |  | |  | |
| Procedure (ordered) Description |  | |  | |
| Order Time |  | |  | |
| Result Time |  | | De-identify dates (Provide year and relative dates e.g. days since 1st encounter); Please include all data from 1/2006 to Present | |
| Result |  | |  | |
| Reference Unit |  | |  | |
| LOINC |  | |  | |
| Other (write in) 2 |  | | Low and High Reference Units | |

Footnotes: **Counts for all lab tests within your cohort can be pulled first for review to allow for selecting the individual test results needed for the study.**

1 Internal CareConnect ID – This is the Ordered Test/Panel Id (e.g. CBC, BMP, Hepatic Function).

2 Other possible variables include: Low and High Reference Units etc.

**Medication Orders (Prescriptions) or Med Administration**

Data is readily available from March 2, 2013. Data prior to this date may be requested but would only cover a minority of patients and will take longer to extract.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Include Variable | | Indicate or describe values to include | | |
| Study ID/Encounter ID | |  | | |  |
| Order Med ID1 | |  | |  | |
| Order Date | |  | | De-identify dates (Provide year and relative dates e.g. days since 1st encounter); Please include all data from 3/2013 to Present | |
| Start Date | |  | |  | |
| End Date | |  | |  | |
| Epic Medication ID | |  | |  | |
| Epic Medication Name | |  | |  | |
| MediSpan Generic Name | |  | |  | |
| MediSpan Class Name | |  | |  | |
| Quantity | |  | |  | |
| Refills | |  | |  | |
| Sig (Dose) | |  | |  | |
| Taken Time (Inpatient Only) | |  | |  | |
| Frequency | |  | |  | |
| Other (write in) | |  | |  | |

Footnotes:

1 Ordering mode may be marked as “Inpatient” if the drug was facility-administered (even if administered at an outpatient facility). Ordering mode may also be marked as “Outpatient” if the prescription is generated for fulfillment at a pharmacy, even if ordered during an inpatient encounter.

**Social History**

Data is readily available from March 2, 2013. Data prior to this date is unlikely to be available.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variable | Include Variable | | | | | Indicate or describe values to include |
| Study ID | |  |  | | | |
| Sexually Active | |  | |  | | |
| Female Partner (Yes/No) | |  | |  | | |
| Male Partner (Yes/No) | |  | |  | | |
| Smoking Tobacco User (Smoking Status)1 | |  | |  | | |
| Tobacco Pack per Day | |  | |  | | |
| Tobacco Used Years | |  | |  | | |
| Tobacco User 2 | |  | |  | | |
| Cigarettes, Pipe, Cigar, Chew, Snuff (Yes/No) | |  | |  | | |
| Smoking Start Date | |  | | De-identify dates (Provide year and relative dates e.g. days since 1st encounter); Please include all data from 3/2013 to Present | | |
| Smoking Quit Date | |  | |  | | |
| Alcohol User3 | |  | |  | | |
| Alcohol Ounces per Week | |  | |  | | |
| Alcohol Comments | |  | | |  | |
| Alcohol Type | |  | | |  | |
| IV Drug User (Yes/No) | |  | | |  | |
| Illicit Drug Frequency | |  | | |  | |
| Illicit Drug Comment | |  | | |  | |
| Other (write in)4 | |  | | |  | |

Footnotes:

1 Values for Smoking Tobacco User are “Current Every Day Smoker”, “Current Some Day Smoker”, “Smoker, Current Status Unknown”, “Former Smoker”, “Never Smoker”, “Never Assessed”, “Passive Smoke Exposure - Never Smoker”, “Unknown If Ever Smoked”, “Heavy Tobacco Smoker”, and “Light Tobacco Smoker”. A patient is considered a current smoker if this variable is one of the following: “Current Every Day Smoker”, “Current Some Day Smoker”, “Heavy Tobacco Smoker”, or “Light Tobacco Smoker”.

2 Yes/no variable based on any of the individual tobacco forms being “Yes.” (Cigarettes, Pipe, Cigar, Chew, Snuff)

3 A patient is considered a current alcohol user if Alcohol User is “Yes” or Alcohol Ounces per Week is greater than zero. (This is our construct, not a variable in Clarity.)

4 Other possible variables include contraceptive types, e.g. condom (yes /no), pill (yes/no), and diaphragm (yes/no), etc., and tobacco forms, e.g. snuff (yes/no), cigar (yes/no), chew (yes/no), pipe (yes/no), etc.

**Family History**

Data is readily available from March 2, 2013. Data prior to this date is unlikely to be available.

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | | Include Variable | Indicate or describe values to include |
| Study ID |  | |  |
| Medical History (Dx) | |  |  |
| Relation | |  |  |

**Allergies**

Data is readily available from March 2, 2013. Data prior to this date is unlikely to be available.

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | | Include Variable | Indicate or describe values to include |
| Study ID |  | |  |
| Allergen | |  |  |
| Description | |  |  |
| Reaction | |  |  |
| Date Noted | |  | De-identify dates (Provide year and relative dates e.g. days since 1st encounter); Please include all data from 3/2013 to Present |
| Allergy Status | |  |  |
| Severity | |  |  |
| Other (write in)1 | |  |  |

Footnotes:

1 Details for allergies may need further investigation.

**Provider Notes (Clinical Documents)**

Data is readily available from March 2, 2013. Data prior to this date may be requested but will take longer to extract.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Include Variable | | Indicate or describe values to include | |
| Study ID/Encounter ID | |  | |  |
| Note Type |  | |  | |
| Note ID |  | |  | |
| Create Date |  | |  | |
| Create By |  | |  | |
| Line # |  | |  | |
| Note Text1 |  | |  | |
| Contact Date2 |  | |  | |
| Other (write in) |  | |  | |

Footnotes: **Counts of Note Types can be pulled first for review and to allow for more detailed selection.**

**Note Text can potentially be searched for specific terms.**

1 The text of pre-Beaker pathology reports (from prior to 3/2016) can be searched in PowerPath with the help of a pathologist collaborator, at $187.50/hour. The search criteria would need to be implemented separately using data available in PowerPath, which includes a specimen code and ICD-9 codes but not SNOMED, surgical procedure codes or most other EHR data.

2 Contact Date is the date the text was entered.

**Pathology & Cytology**

Data is readily available from March 2, 2013. Data prior to this date may be requested but will take longer to extract.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Include Variable | | Indicate or describe values to include | |
| Study ID/Encounter ID | |  | |  |
| Accession Number |  | |  | |
| Order Procedure Id |  | |  | |
| Order Time |  | |  | |
| Result Time |  | |  | |
| Description |  | |  | |
| Component Name |  | |  | |
| Component Comment |  | |  | |
| Line # |  | |  | |
| Line Comment |  | |  | |
| Results Comment (Narrative) |  | |  | |
| Specimen Source |  | |  | |
| Specimen Type |  | |  | |
| Specimen Taken Time |  | |  | |
| Component Result |  | |  | |

Footnotes: **Counts of orders by Procedure Name can be pulled first for review and to allow for more detailed selection. Result Text can potentially be searched for specific terms.**

**Imaging Orders and Results**

Data is readily available from March 2, 2013. Data prior to this date may be requested but will take longer to extract.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Include Variable | | Indicate or describe values to include | |
| Study ID/Encounter ID | |  | |  |
| Order Procedure ID |  | |  | |
| Order Time |  | |  | |
| Result Time |  | |  | |
| Procedure Code |  | |  | |
| Procedure Name |  | |  | |
| Accession Number |  | |  | |
| Narrative |  | |  | |
| Impression |  | |  | |

Footnotes: **Counts of orders by Type can be pulled first for review and to allow for more detailed selection.**

**Result Text can potentially be searched for specific terms or string patterns (regular expressions).**

**Narrative and Impression will each be sent as a separate table with order procedure id, line, and text fields.**

**Cultures/Isolates (Microbiology)**

Data is available from 2006. Data prior to this date may be requested but will take longer to extract.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Include Variable | | | Indicate or describe values to include / Comments | |
| Study ID/Encounter ID | |  | | |  |
| Order Procedure ID |  | |  | | |
| Order Time |  | |  | | |
| Result Time |  | |  | | |
| Procedure Code |  | |  | | |
| Procedure Name |  | |  | | |
| Specimen Source |  | |  | | |
| Specimen Type |  | |  | | |
| Line |  | | | Used to sort results | |
| Result |  | | |  | |
| Component |  | | |  | |
| Component Comment |  | | |  | |
| Organism Name |  | | |  | |
| Line Comment |  | | | Used to sort results | |
| Results Comment |  | | |  | |

Footnotes: **Counts of orders by Type can be pulled first for review and to allow for more detailed selection.**

**Results prior to 3/1/2013 will have similar layout to the pathology & cytology table above.**

**Cultures/Isolates Susceptibility & Sensitivity (Microbiology)**

This section will be pulled based on the main microbiology extract above. Because there are multiple entries per organism, these are sent in a separate table. All fields below will be pulled when “Include Data” box is checked.

|  |  |  |
| --- | --- | --- |
| Variable | Include Variable | Comments |
| Order Procedure ID |  | This field will link to the micro order in the table above |
| Organism Name |  |  |
| Susceptibility |  |  |
| Sensitivity |  |  |
| Antibiotic |  |  |