

C1 Activity Methods [DRAFT]

January 3, 2016

1 Data Sources

Multiple datasets were combined to address questions in the current investigation.

1. A **PHOCIS Referral** record (Form #473) represents one referral received by a C1 program, which roughly corresponds to one record per mother. A record was removed if it's 'ReferralDate' fell outside 2006-01-01 and 2015-01-01 (recall ETO replaced PHOCIS as the primary C1 data system January 2015). When duplicate records existing for a parent, only the most recent was retained, which reduced the count roughly from 47,000 to 43,000 records. See below for a discussion of 'Suitability'.
2. A **PHOCIS Client Activity Status** record (Form #439) represents a change in a mother's C1 status; mothers typically have multiple records in this dataset. See below for a discussion of 'Suitability'.
3. A **PHOCIS Home Visit Encounter** record (Form #407) represents any completed or failed visit attempt.
4. An **ETO Referral** record (a combination of the Demographics and other TouchPoints) represents one referral received by a C1 program. Unlike the PHOCIS Referral source, the ETO Referral source did not contain items determining suitability.
5. An **ETO Screening** record (Screening TouchPoint) represents one potential client received by a C1 program, and is linked to a referral record through the client's Case Number. This TouchPoint contained RF20, RF22, and RF23 found in the PHOCIS referral form, which could determine suitability.
6. An **ETO Enroll** record (Enrollment TouchPoint) represents a change in a mother's enrollment status. For the purposes of our current analyses, this TouchPoint resembles the PHOCIS Client Activity Status form. The status of a client's first record is typically 'pending'. The second is typically either 'rejected' or 'accepted'.
7. An **ETO Visit** record (Encounter TouchPoint) represents any completed or failed visit attempt.
8. The OSDH **T&E** (Time and Effort) database contains the amount of time each month each OSDH employee dedicated to C1 (and other OSDH programs). The activity is further categorized by type (eg, administrative leave, program support). Each record is a combination of employee by month.
9. A **WIC Estimate of Need** record represents the number of women, infants, and children in need for a particular year. The county's number of infants in need was used in these analyses. The estimates are available for 1998, 2000, 2004, 2005, 2006, 2010, and 2014. Loess regression was used to smooth existing years and approximate the years lacking an estimate. The dataset and further details are available on our project's website. The count of infants in need was helpful in many analyses to make comparisons between counties more comparable. For instance, we typically preferred examining trends in 'referrals per need' than trends in 'referrals counts'.
10. A **Program** record represents a physical C1 site. These 223 records contained variables such as the clinic's county, and were combined from the PHOCIS and ETO systems.
11. A **County** record represents one of the state's 77 counties. It contained variables such as the county's Lead Nurse Region.
12. A **Region** record represents one of the state's C1 Lead Nurse Regions.

13. A **PHOCIS CodeTable** record translates numeric lookup codes into labels meaningful to humans. Each of the 1,337 records was a possible response, although only a few dozen were required for this project.
14. The ETO system lacked an explicit CodeTable, and exported choices only as text. The possible lookup values were inferred from the response collection. For many multiple choice items, a secondary lookup table was created to collapse multiple values into the same response. For example, “All nurses have full caseloads” and “All staff have full caseloads” were different options available at different times during ETO’s deployment, and were mapped to the same response in the analysis. Similarly, “She won’t be raising the child” and “She wona??t be raising the child” were considered equivalent.

2 Milestones

2.1 Suitable Referral

A potential client passes through several steps before completing a referral. Slight differences in the referral process existed (a) between the PHOCIS and ETO data systems, (b) over time within the data systems, and (c) between clinics (and perhaps nurses). Some procedural differences were unintentional, while others were adapted to best meet the local needs; for example, the client pool and environment is quite different (and deserves a different approach) between rural and urban regions. In this investigation, ‘Referral count’ was operationalized to best accommodate the 20 Oklahoma regions from 2009 through 2015.

‘Suitability’ is defined similarly to the official C1 ‘eligibility’, but is more inclusive. A single referral’s suitability could be informed by the Referral and the Enrollment data sources. The three referral items were “Was an interview ever completed with the referred individual about participating in Children First?” (RF20), “Was the initial home visit completed?” (RF22), and “Did the referred individual ever want to participate in the Children First program?” (RF23). The client’s enrollment reflected either the Client Activity Status form (in PHOCIS) or the Enrollment TouchPoint (in ETO). If any of the four referral or enrollment responses was ‘no’, the nurses chose from 18 possible reasons. Table zzz-suitability-classification-zzz enumerates the 18 reason codes, and how each was classified as either ‘suitable’, ‘not suitable’, or ‘no contact’.

A single suitability value was assigned to each referral by combining the three referral responses and the most recent enrollment value. Suitability was determined by RF20 or RF22 if either value was present. If these were missing, then the enrollment item was used. If these were missing, then RF23 was used if it was present. Sometimes multiple (and conflicting) values existed for a referral. A nonsuitable value trumped any suitable value; however any nonsuitable response was trumped by a successful enrollment (and the referral was considered suitable).

For PHOCIS data, a few extra steps were taken to align client data between the referral form and client activity status form. The C1 program currently requires all potential clients have at least one referral record; an Activity Status record is required only if zzz {Question for Miriam or John}. However the operating practice for some counties during some years relaxed this requirement and consequently 15% of mothers present in the Activity Status dataset did not have a Referral record; however only 5% of mothers present in the Referral dataset did not have a Activity Status record. An outer join created an overall list containing all unique members in both datasets. The mother’s earliest date (found in either dataset) was important in later longitudinal analyses because it indicates when the mother first interacted with C1, while avoiding the holes in the two forms’ dates. Although both forms had a field for the county, the form’s clinic variable was used to determine the county, since it was entered much more reliably.

2.2 Completed Enrollment

The definition of a successful enrollment was constant across time and programs. Our operationalization matched the official C1 protocol: the client must complete at least one visit. The two relevant data sources were the PHOCIS Home Visit Encounter form and the ETO Enounter TouchPoint.

2.3 Completed Visit

The operationalization of a completed visit was straight-forward too. The PHOCIS and ETO visit data sources explicitly indicate if a visit was completed successfully.

2.4 Graduation

Although the C1 protocol defines a client’s graduation (and PHOCIS and ETO contained explicit graduation fields) it was not used. C1 personnel advised us that these variables are not completed consistently, and our preliminary examination supported this. Instead, graduation was operationalized as the completion of 40 or more visits.

3 Combining PHOCIS and ETO sources

C1’s live data system transitioned from PHOCIS and ETO on January 1, 2015, and accordingly, many PHOCIS records from 2013 and 2014 were copied to ETO. When a record existed in both systems, the ETO record took precedence. Deduplication was fairly simple, because most ETO records contained PHOCIS’s unique client ID. Referral records and visit records were considered duplicates if the client ID and date matched.

4 Estimating Nurse Days

Personnel information was combined with referral and visit data to assess a region’s time- and cost-efficiency. Several approaches were used during the investigation’s life; the eventual approach used only the OSDH ‘Time and Effort’ (T&E) data system, whose first record was June 2012; recall the analyses’ client information starts January 2009. One record corresponds to one month for a C1 employee. The record contained the given month’s FTE, and the employee’s C1 start date and OSDH start date. Information about leave and training was sparse and unreliably represented, and inconsequential to later analyses.

One ‘Nurse Work Day’ represents every 8 hours that a C1 nurse is available to see clients, and this estimation involved several calculations. The FTE was multiplied by either 0, 0.5, or 1.0, depending on the nurse’s C1 tenure. Lead nurses, and nurses hired within 2 months are not required to make any visits; nurses in their third through ninth month carry a half-load. This was multiplied by the number of weekdays in the month, minus the estimated leave. Because specific leave information was not available in the T&E system, we approximated the leave for each 1.0 FTE nurse at 27 days a year (we approximated 8 of annual leave, 8 of sick leave, and 11 of holiday leave).

Monthly sums were calculated for each county. During the system’s first 18 months (between July 2012 and December 2013), 6 months were missing. The missing county sums were linearly interpolated. We were comfortable with this because the FTE sums were (a) fairly stable and (b) not used as the primary outcome. Their main purpose was to scale outcomes like referral and visit counts. (This data manipulation inspired a recent workshop that used synthetic data).

NOT USED

1. Overview: This primary question addressed by this report is, “How efficiently and thoroughly was the state covered by Children First program (C1) for 2009 through 2015?” The 20 regions in Oklahoma (officially called “Lead Nurse Regions” within OSDH) serve as the unit of analysis. A client’s progress through the program was assessed by their (a) referral, (b) enrollment, and (c) visits. There are at least three perspectives of the primary question. For each perspective, the longitudinal trends and between-region variability received extra attention.
 - (a) How many C1 clients were served?
 - (b) What proportion of infants in need were served (where need is estimated by WIC)?
 - (c) How many referrals were obtained?
 - (d) How many referrals were successfully converted to enrollments and initial visits?
2. Determine Region Membership of Nurse: The T&E database was the primary sourced used to determine the amount of OSDH funds spent in each region. However, the T&E records did not indicate which clinic, county, or region benefited from the nurses’ time. [The patient records, which had the nurse’s ID, were used to determine the nurse’s mode for each month.]
3. A dataset of the **C1Nurses**, compiled for us by OSDH. This links the nurse names (in the T&E dataset) to the nurse IDs (in the PHOCIS datasets). Each of the 453 records is a current or former C1 nurse in the [rural counties].