Standard Operating Procedures for Water and Addresses

Basic guidelines for selecting maternal addresses and associated relevant water nitrate values

4/14/2009

Prepared by Griesenbeck JS, Brender JD, Weyer PJ, Horel SA

1. SOP for Water and Addresses

Each ID is listed with the addresses (residence sequence)they have lived at, the dates they lived at those addresses, and the dates surrounding their conception (including 1 month prior to conception (B1), 1st month of pregnancy (P1), 2nd month of pregnancy (P2), and 3rd month of pregnancy(P3). These variables were created internally using a STATA program developed by Dr. Chuck Huber.

If a person lived at an address during any of the designated months surrounding conception, code that time period as" yes"--in our case "1"-- (ResB1, ResP1, ResP2, and/or ResP3). If the address does not correspond with the time period code it as "no"—in our case "2". So 1= yes and 2= no.

2. Assumptions, definitions, and helpful hints:

1 month equals 30 days.

Start month is defined as the first day of the month listed as beginning the period of residence at each address because the address database only lists month and year but not day.

Stop month is defined as the last day of the month listed as ending the period of residence at each address because the address database only lists month and year but not day.

To receive credit for having lived at a residence for that month, the mother must have lived at that residence for at least 15 days out of the 30 day month.

Months coded as "98" under the start or finish month categories are considered as midpoint of the calendar year (June 15, XXXX).

It is possible to have overlapping residencies and more than one residency representing a specific period of time.

When the start month or end month is missing, the values for residency are coded as "unknown or missing" and we code those as "9".

Values for start and end dates for "year" occasionally have 9998 or other values, these observations are considered individually and all information is taken into account prior to determining the residency code for that mother. There should be relatively few of these.

We found that it is easier to consider each line of data individually when coding for residency qualification, rather than the total number of residencies simultaneously; this can help reduce assumptions about residency status and potentially reduce clerical errors as well.

The data was first hand coded and then verified with a systematic coding schema using STATA. Quality control checks were performed on every 10^{th} observation and every observation where manual coding and computer coding assignments differed.

3. An example on how to code using 1 as "yes" and 2 as "no":

Case# XXXXXXXXX lived at 2 addresses

She lived at address 1 from 5/1999-9/1999

She lived at address 2 from 9/1999-2/2000

Her conception information is as follows:

			Address1	Address2
StartB1	13/Oct/1999	ResB1	2	1
EndP1	12/Dec/1999	ResP1	2	1
EndP2	11/Jan/2000	ResP2	2	1
EndP3	10/Feb/2000	ResP3	2	1

When looking at the conception information (B1, P1, P2, P3) look at is as the entire month to determine if the address falls on the same time period, and be sure to make sure the address has at least a 15 day overlap with the pregnancy month.

Separate example: If the address went to May 1999 and EndP3 was June 8, 1999, ResP3 would still be coded as "yes" for that residence because the majority of the month falls in the same time period and less than 15 days are accounted for in June.

3. How to link addresses to specific water data

A. If the address is linked to a water district:

If the address is linked to particular water district (has a CCN and name of district associated with the address of interest), then the time period associated with that district should be assessed (based on month and year of residence). For the period B1-P3, only those months that are coded as "yes" (the mother lived at that address during that time frame) will be used to identify relevant nitrate values. If B1, P1, P2, and P3 are all coded as "yes", then the time period from B1-P3 is the relevant time period. If only P1-P3 are relevant, then it is important to remember that P1 is the end of the period so one month prior to the P1 date

through the P3 date would be the relevant period. Likewise, if only B1 is relevant, then it is important to remember that B1 is the beginning of the period so B1+ 1 month becomes the relevant period, and so on.

1. Selecting relevant nitrate values:

Once the time period of interest is determined it is necessary to select which nitrate value(s) best represents the exposure. Using the access database provided by TCEQ, we selected only those values that reported Nitrate values --Nitrate as Nitrogen-- and treated water. Flowcharts 1-4 may be used as a visual representation of how nitrate values are selected. The algorithm for selecting values is as follows:

1. If there is a value from the water district associated with the address during the actual dates of residence that falls within B1-P3, then that value (or values, if more than one) will be entered into the database.

2. If no values are reported during the actual dates of residence, then one year prior to the earliest relevant date through one year after the last day of residence at the qualifying address will be assessed for values. All values reported during this "one year before through one year after" window will be included in the database.

3. If no values exist for steps 1 or 2, then the value located closest to the earliest date of residence and the value reported which is closest to the last day of residence will be entered into the database. If there either the pre or post value is missing, simply use the one that is available.

4. In some rare cases, there may be more than 20 nitrate values reported during a relevant period. In those cases, average the nitrate values reported by month and report the test date as the first day of the month under the "DateNit1" column. Enter the average values for that month under the "Nitrate1" column and enter the range under the "DetLim1" column. For the next month with multiple values repeat the procedure but report the values as described in this paragraph under "DateNit2", "Nitrate2", and "DetLim2" as appropriate. Repeat the procedure as necessary. Ensure that the "DetLim" fields are coded as string variables to allow you to enter the range of values for that month (minimum and maximum). More information regarding the specific fields is available in section 2 below.

2. Entering nitrate values into the database:

For each nitrate value selected for entry into the database, other information must also be entered including the date of the value and if the value is below the detection limit.

Every nitrate value should have a field such as DateNit1 (Date of the first Nitrate Value), Nitrate1 (the actual nitrate value for the first relevant sample in mg/L), and DetLim(code this as yes if the value entered is half of the reported detection limit). As for values reported that are <.01 mg/L (or any value reported as "less than") one-half of the value will be entered as the nitrate value for that entry. For example, if a relevant value is reported as <.01, then .005 should be entered under the Nitrate1 value (if it is the first, Nitrate 2 if the second, Nitrate3 if the third, etc). Additionally, if a value is entered that is ½ of the detection limit, then DetLim1 (or DetLim2, or DetLim3, etc) should also have a value (we use "1" in SPSS) that denotes the value for that entry is one that was reported below the detection limit.

B. If the address is not linked to CCN, but is linked to City Name

Look for relevant values for that city in the TCEQ database. In some cases it is very straight forward, and there is only one water provider for that city. However, when multiple water providers are listed for a city, the municipal water provider that supplies most of the city is used.

If city is not listed in TCEQ, then contact the city water district directly for nitrate values. And if the city can not be contacted by phone/email, then attempt to locate water reports on-line.

C. In cases when zip code and city are inconsistent

Make a decision to use city and state, regardless of zip code or street address.

D. In cases where there are multiple PWS's for CCN

The CCN will be contacted for a map (shape file) of the PWS boundaries within their CCN. The PWS boundaries can be used to link the particular PWS to the qualifying residence.

E. In cases where CCN or city is not listed in TCEQ

When contacting water supplier, first contact by phone and email. If they do not respond, then contact with an official letter. If no response after one month, then call to follow up or answer any questions. If no response after two more weeks then use substitute values from TCEQ and if city not in TCEQ then search for the most appropriate values on-line.

F. In cases where the participant moved out of state

When searching for an out of state water value, first search the city or town's website. Most cities have water quality reports posted on either their city website or their water supplier's website. If the website has multiple years of water quality reports, select the date that is closest to when the participant lived in the city. If multiple water supply centers are used for one city, use the location that is closest to the zip code listed in the database.

If the city or town does not show their water quality reports online, search for the water supplier or utility company that provides for the city/town. If it is unclear who provides water to the city/town, contact the city or town hall. The operator at the town hall often has the water quality reports, or will be able to direct you to the local water service provider.

After finding the nitrate value for the out of state location, enter it in the database along with the range of values if they exist. If only a range is given on the water quality report, take the median between the upper and lower values detected, and note the range in DetLim1.

a. Example on how to enter selected values:

Case# XXXXXXXXX lived at 2 addresses

She lived at address 1 from 5/1999-10/1999

She lived at address 2 from 11/1999-2/2000

Her conception information is as follows:

			Address1	Address2
StartB1	13/Oct/1999	ResB1	1	2
EndP1	12/Dec/1999	ResP1	2	1
EndP2	11/Jan/2000	ResP2	2	1
EndP3	10/Feb/2000	ResP3	2	1

Address	**StartB1**	**EndP1**	**EndP2**	**EndP3**	**CCN**	**Water District**	**Date Nite1**	**Nitrate1**	**Det lim1**	**Date Nite2**	**Nitrate2**	**Det lim2**	**Date Nite3**	**Nitrate3**	**Det lim3**
1: 5/99-10/99	1	2	2	2	60512	Port Lavaca									
2: 11/99-2/00	2	1	1	1	52318	BMWD									

Note: Using this information we know that for **address 1** the period of interest is **Oct/99-Nov/99**

For **address 2** the period of interest is **Nov/99-Feb/00** (the period starts in Nov, because P1 ended in December and we have to account for the entire month)

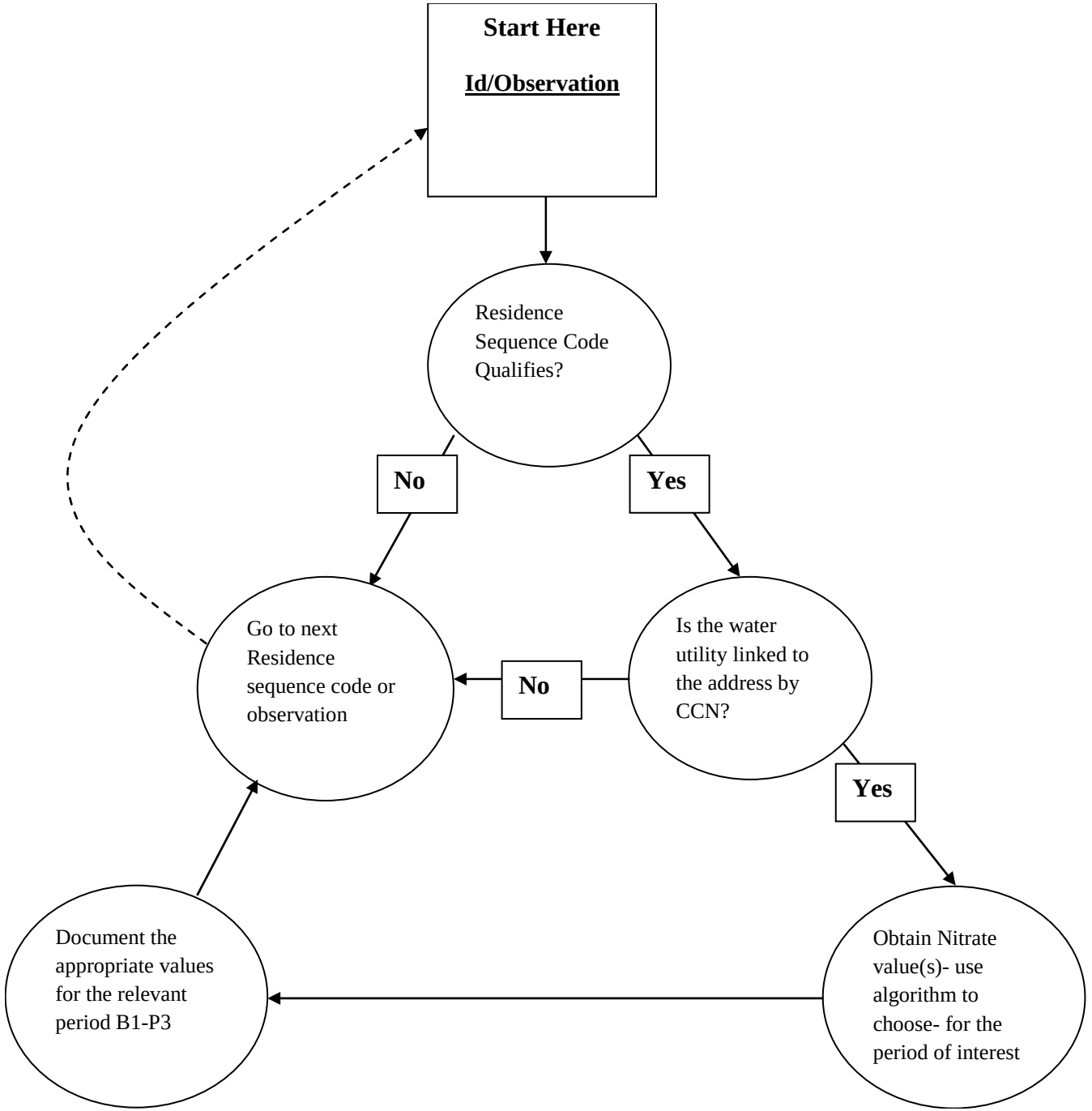
Here is an excerpt from the TCEQ water database:

CCN	Water Dist	Sample Date	Test Name	Result	Comments
60512	Port Lavaca	**10/21/99**	**Nitrate**	**1.7**	This value falls within the period (use)
60512	Port Lavaca	**10/30/99**	**Nitrate**	**<.01**	Use .005 as half the detection limit (use ½ det limit)
60512	Port Lavaca	12/2/99	Nitrate	2.0	This value does not fall within the period (don't use)
52318	BMWD	7/20/98	Nitrate	0.5	Neither in the dates of residence or in the year before or after (don't use)
52318	BMWD	**8/12/99**	**Nitrate**	**1.5**	Not in actual dates of residence but does fall within the one year prior/post window (use)
52318	BMWD	**5/6/00**	**Nitrate**	**1.8**	Not in actual dates of residence but does fall within the one year prior/post window (use)
52318	BMWD	**2/10/01**	**Nitrate**	**0.9**	Not in actual dates of residence but does fall within the one year prior/post window (use)

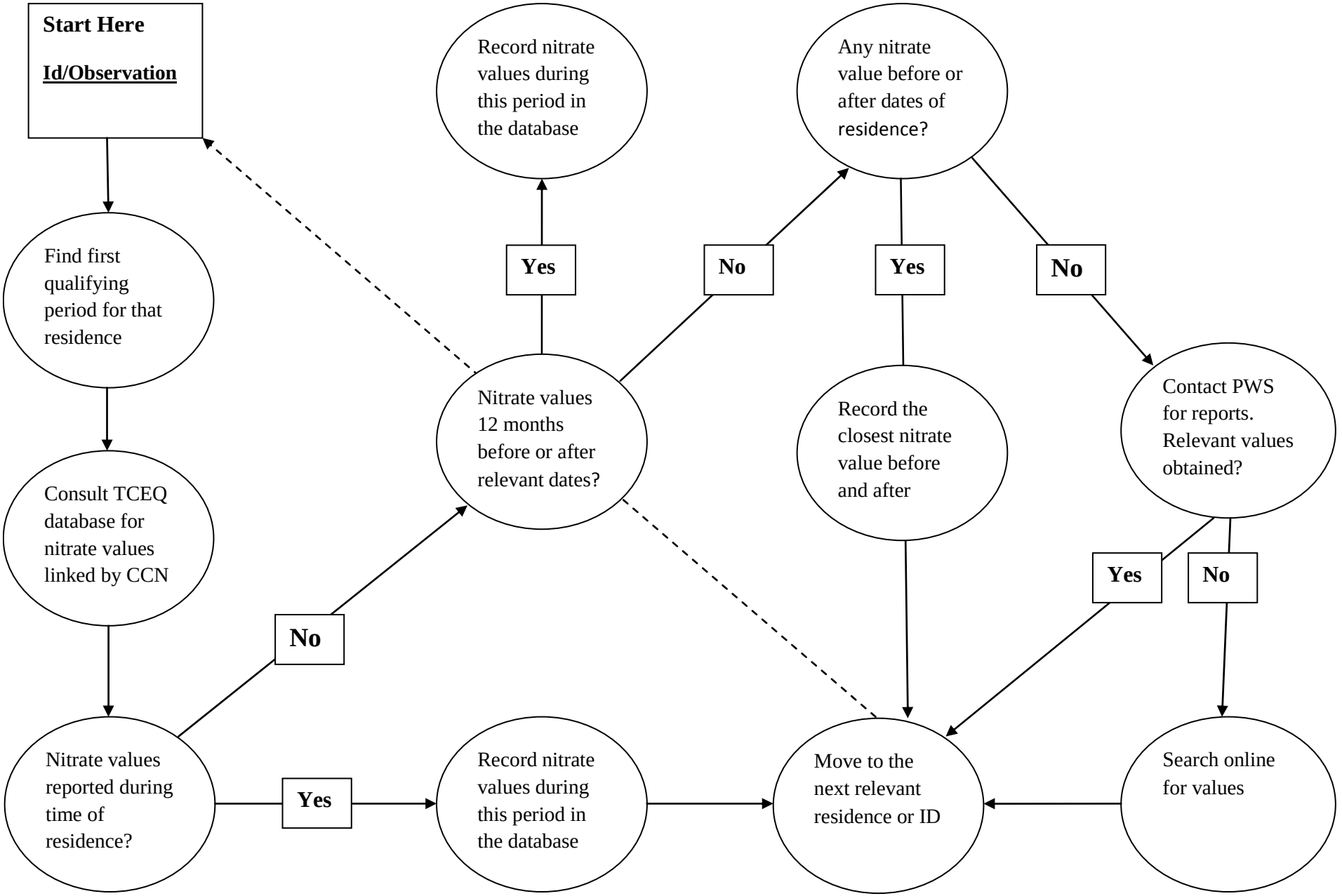
The table would be filled in as follows:

Address	StartB1	EndP1	EndP2	EndP3	CCN	Water District	DateNite 1	Nitrate1	Det lim1	DateN ite2	Nitrate2	Det lim2	Date Nite3	Nitrate3	Det lim3
1: 5/99-10/99	1	2	2	2	60512	Port Lavaca	10/21/99	1.7		10/30/ 99	.005	1			
2: 11/99-2/00	2	1	1	1	52318	BMWD	8/12/99	1.5		5/6/00	1.8		2/10/0 1	0.9	

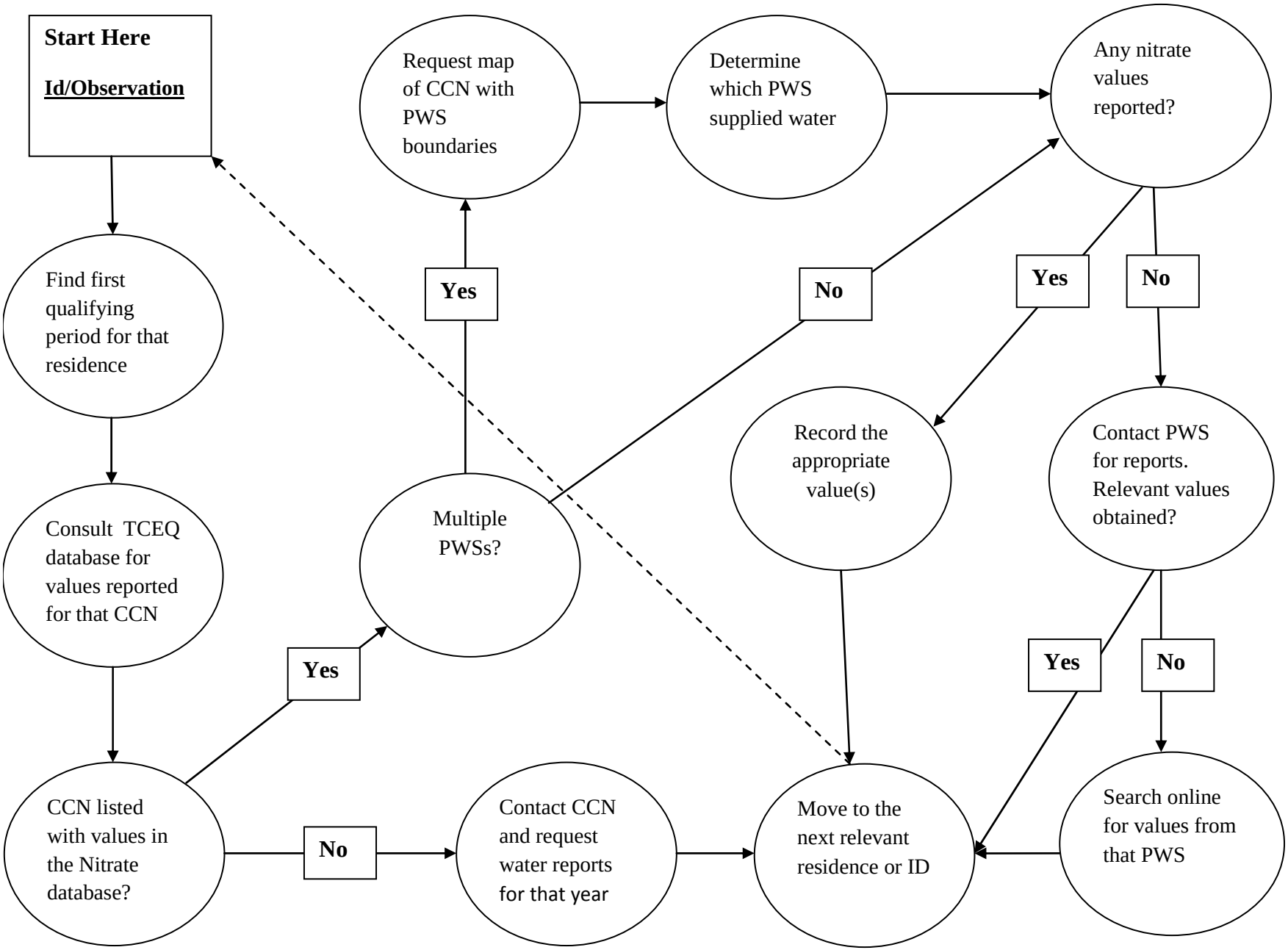
Flowchart 1: Overview of the process for selecting and assigning nitrate water values to maternal residences



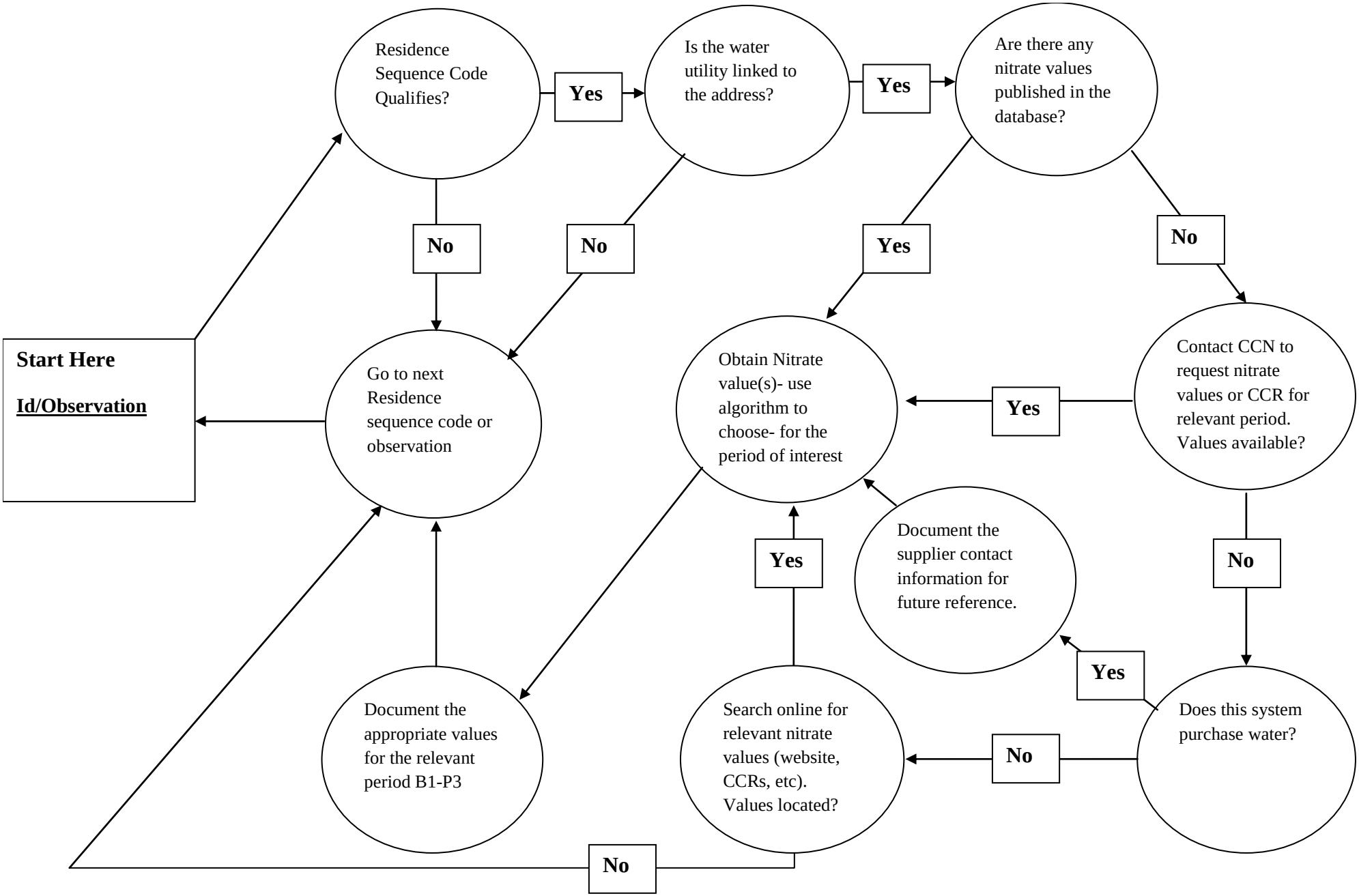
Flowchart 2: Process for selecting relevant nitrate values with addresses linked to CCN



Flowchart 3: Process for selecting and locating nitrate values for multiple PWS within a single CCN



Flowchart 4: Process for assigning nitrate values to maternal residences when no values are reported for an address linked to a CCN



4. Calculating Nitrate Values for B1-P3

To calculate nitrate values for each participant, take the average of all nitrate values listed as previously discussed. If a participant has lived at more than one address during the periconceptional period, then the nitrate value for that participant will be calculated as follows:

a. First calculate the average nitrate amount for each address.

b. Then, weight the address according to how long the participant lived at the address. Weighting scale is determined based on how many months they lived at that address out of the 4 month periconceptional period.

   For Example: A participant lived at two addresses during their periconceptional period. They lived at the first address for the month before conception (B1-Conception Date) and the second address from conception to 3 months after (Conception Date-P3). The first address would weight 25% (1 out of 4 months) and the second address would weight 75% (3 out of 4 months).

c. Each weight is then multiplied by the average nitrate value calculated for that address. All of the values are then added together to get the total nitrate level for that participant.

   If a participant showed to live at 2 addresses during the same month, they were considered to live at each address for half of a month. The weighting scale was adjusted accordingly.

5. Calculating Nitrate Values for B1-P1

To calculate the values for B1-P1, most of the same procedures were used as above. For each participant, the nitrate value was calculated by looking at which sampling dates fell within the time period listed for B1-P1. If no samples were taken during the dates listed for B1-P1, then the sample date listed closest to the B1-P1 date's value was used. In the case where a participant lived at more than one address for the B1-P1 period, the same weighting method was used above for the 2 month time period.