Comparing CAPI Trace File Data and Quality Control Reinterview Data as Methods of Maintaining Data Quality

Matt Jans[1], Robyn Sirkis[2], Christina Schultheis[3]
Renee Gindi[4], James Dahlhamer[4]
[1]U.S. Census Bureau, 4600 Silver Hill Rd., Washington, DC 20233
[2] U.S. Census Bureau, 4600 Silver Hill Rd., Washington, DC 20233
[3] U.S. Census Bureau, 4600 Silver Hill Rd., Washington, DC 20233
[4]National Center for Health Statistics, 3311 Toledo Rd., Hyattsville, MD 20782
[5] National Center for Health Statistics, 3311 Toledo Rd., Hyattsville, MD 20782

Abstract
Quality control is paramount to surveys for which the U.S. Census Bureau is the data collection agent. Two methods currently used to measure data quality are the Performance and Data Analysis (PANDA) system and Quality Control (QC) Reinterview. PANDA uses CAPI trace files, data files, and other case information (e.g., interview date and time) as indicators of cases or interviewers that might be at risk for lowering overall data quality (e.g., overnight interviews as indicators of falsification). The QC Reinterview is a verification interview with respondents that asks questions about the interview experience to detect falsification. This paper explores whether these systems capture the same cases and interviewers. Both the PANDA and QC Reinterview are used on the National Health Interview Survey, a nationwide face-to-face CAPI survey sponsored by the National Center for Health Statistics. Using data quality results from 2009 and 2010, we analyze which cases (i.e., sample units or households) and interviewers are identified by PANDA, by QC Reinterview, or by both. We use a multinomial logistic regression predicting identification by PANDA, Reinterview, both, or neither to see if any sample or caseload factors predict identification. Our results suggest independence of the systems with some qualifications. We discuss avenues for further research.

Key Words: paradata, quality control, field methods

1. Introduction

Quality control is an important part of any survey data collection operation. Quality assurance (QA)[1] techniques, like reinterviewing respondents or using systematic procedures for identifying interviewers who produce outlier data on key performance criteria (e.g., item nonresponse rates and interview durations), can be used to catch errors in data collection early, allowing for correction before

[1] We use the phrases "quality control" and "quality assurance" interchangeably in this paper. The systems reviewed in this paper are often thought of as "quality assurance" systems, but the use of paradata on which we report has features of quality control not completely different from statistical process control. For an in-depth discussion of statistical process control opportunities with the NHIS, see Sirkis & Jans (forthcoming) in the Proceedings of the 2011 Joint Statistical Meetings.

delivering data to a client. Yet any production operation adds additional cost and time to survey budgets, so an implementation decision should be based on the net benefit of a given program. Benefit can be gauged by numerous metrics, one of which is how many unique problem cases the QA technique detects before data are delivered.

In this paper, we explore the potential overlap in two QA systems currently in use with the National Health Interview Survey (NHIS): the Performance and Data Analysis (PANDA) system and Quality Control (QC) Reinterview. More specifically, we address the following research questions:

1. Do the two QA systems identify the same problems?

2. What predicts identification of problems in either or both systems?

Highly redundant systems would be inefficient, especially in the current climate of shrinking funding for surveys. Yet two systems with unique contributions to quality control expand the range of confidence we can have in our data. In the remaining sections we describe the NHIS and the two QA systems currently in use (section 1), the data and analytic methods used to compare and contrast the two systems (section 2), the results of our analysis (section 3), and conclude with a discussion of limitations and future steps in evaluating QA systems (sections 4 and 5).

1.1 The National Health Interview Survey

The National Health Interview is an annual survey of the health of the civilian, noninstitutionalized household population of the United States, and is conducted by the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC). The survey utilizes a multi-stage, clustered sample design, with oversampling of black, Hispanic, and Asian persons, and produces nationally representative data on health insurance coverage, health care access and utilization, health status, health behaviors, and other health-related topics. The microdata are released on an annual basis, approximately six months after the end of the data collection year.

Roughly 600-700 U. S. Census Bureau interviewers conduct the in-person interviews (some telephone follow-up is allowed) using computer assisted personal interviewing (CAPI). In 2009 and 2010, the data years included in this analysis, cases were assigned to interviewers each week throughout the calendar year (except the first two weeks of January) and were to be completed within a 17-day interview period. Each year, interviews are conducted in approximately 35,000 households yielding data on roughly 87,500 persons.

The core survey instrument contains four main modules: household composition, family, sample child, and sample adult. A household respondent provides demographic information on all members of the household in the household

composition module. For each family within a household, the family module is completed by one family respondent who provides sociodemographic and health information on all members of the family. Additional health information is collected from one randomly selected adult (the "sample adult") aged 18 years or over, and from the parent or guardian of one randomly selected child (the "sample child") under age 18 (if there are children in the family).

1.2 Quality Assurance in the National Health Interview Survey

As noted previously, there are two QA systems currently used with the National Health Interview Survey (NHIS): the Performance and Data Analysis (PANDA) system and Quality Control Reinterview. Each has a unique history and quality assurance goal.

1.2.1 PANDA

The Performance and Data Analysis system (PANDA) is a Web browser-based tool that summarizes and presents collected survey data as well as paradata captured via interviewer observations and audit trails (i.e., files recording keystrokes captured by the CAPI instrument) that are created every time an interviewer (i.e., field representative or "FR") enters a case (O'Reilly, 2009; Rowe, 2009). It is used to track and manage the progress and performance of Census Bureau FRs during a survey period, and it provides an early indication of the quality of the data being collected. The paradata monitored include item nonresponse rates, case completion rates, interview start dates and times, and interview lengths. This tool provides critical, detailed, and time-sensitive data to Census Headquarters (HQ) and Regional Office (RO) staff to identify FRs experiencing performance issues, such as difficulty with survey items or interviewer protocols, or falsification of data. Real-time review of these paradata assures that action can be taken to correct the problem during data collection, minimizing the impact on data and estimates downstream. Originally developed and implemented for the 2007 American Housing Survey, PANDA was implemented for the NHIS in 2008.

The development and implementation of PANDA on NHIS is a case study in successful collaboration between survey sponsors and data collectors (Taylor, 2009). Frontline paradata review is done by Census Bureau field staff and reported to the survey sponsor approximately every month. Modifications to the PANDA system, including the types of indicators reviewed and their presentation, have been made in collaborative interactions between the Census Bureau's NHIS management staff and the NHIS staff of NCHS. Since PANDA's inception in 2008, changes have been made to criteria for cutoff flags (e.g., changing from upper or lower 10% to upper and lower 1% and 5% flag cutoffs based on interviewer-level distributions), and to the format of reports (e.g., from individual reports for sets of flags to overall flags, and summary reports showing the number of flags). More information about the technical specifications of the PANDA interface and its history can be found in the user manual developed by the U.S. Census Bureau (2009).

NHIS paradata that are viewed in PANDA are organized at the national, RO, FR, and individual case (i.e., household) levels. Summary reports can be viewed at will, and data are updated every week. Most reports on PANDA for the NHIS include flags that identify outlying interviewers and/or cases based on various performance and quality indicators. For example, the FR Level Interview Time Report presents the mean, median, minimum, and maximum durations for complete interviews and main interview modules (family, sample child, sample adult) for each interviewer. Flags for the entire interview and each of the main modules are included. For each, a flag of "1" indicates that an interviewer worked 20 or more complete interviews and her/his median interview or module time fell in the bottom 1% of the distribution of all interviewer median interview or module times. A flag value of "2" for the interviewer would indicate that her/his median interview or module time fell between the bottom 1% and bottom 5% of the distribution of interviewer median times. The report allows managers to quickly identify interviewers with extremely short interview times and alert the necessary staff to take appropriate action. Table 1 shows the indicators we focused on for this study, while Appendix A presents the additional indicators that are monitored in PANDA but that we did not use in this analysis.

Table 1: Indicators Monitored in PANDA that Were Used as "PANDA Problem Flag"

Indicator	*Description*	*Purpose*
Overall Interview Duration	Overall interview durations for complete interviews	NHIS flags short overall interview durations to capture falsification. Shortness is defined by the bottom 1% and 5% of interview durations
Family Interview Duration	Family interview section durations for complete interviews	Same rationale and flagging criterion as overall interview duration
Sample Adult Interview Duration	Sample Adult interview section durations for complete interviews	Same rationale and flagging criterion as overall interview duration
Sample Child Interview Duration	Sample Child interview section durations for complete interviews with a Sample Child	Same rationale and flagging criterion as overall interview duration
Total Family Income Item Nonresponse Rate	Proportion of cases where the total family income item was asked and the respondent answered Don't Know or Refused	NHIS monitors high nonresponse rates in the upper 1% and 5% of total family income nonresponse rates as indicators that FR needs retraining in obtaining this sensitive information

Sample Adult Work Items Nonresponse Rate	Percentage of cases where three Sample Adult work items were asked and respondent answered with Don't Know or Refused	Same rationale and flagging criteria as income item nonresponse rate

1.2.2 Quality Control Reinterview

Quality Control (QC) Reinterviews are another way that the NHIS maintains data quality. The primary goal of the QC Reinterview is to detect FR data falsification, but an intended side effect is that interviewers will not falsify because they know it exists. QC Reinterviews involve conducting a follow-up "interview" with a household to determine if an interview was actually conducted in accordance with NHIS procedures. QC Reinterviews are conducted by telephone, usually within two weeks of the NHIS interview. There are two types of QC Reinterviews used in the NHIS: random and supplemental. In the *random* reinterviews, five to ten percent of the original survey sample is preselected randomly and assigned to QC Reinterview. Reinterviews are designed so that each interviewer has cases reinterviewed at least once a year. In the *supplemental* QC Reinterviews, each survey supervisor at each RO selects interviewers or cases to be reinterviewed. For interviewers, this could be because they have had a problem with their workload, including being flagged in PANDA. These problems can be as severe as, but are not limited to, suspicion of falsification. Supplemental Reinterviews are also often used to check the work of a new interviewer, but may be used for any reason the RO has for reinterviewing the case.

The QC Reinterview instrument is not a full re-asking of the original survey questions, but rather a check that verifies whether certain questions were asked and certain procedures were followed during the original interview. There are eleven possible QC Reinterview discrepancies for any given reinterview. These discrepancies arise because, upon reinterview, it is discovered that specific interviewing procedures were not followed by the interviewer. The presence or absence of QC Reinterview discrepancies aids the reinterviewer in determining whether or not falsification occurred.

In 2009, the NHIS QC Reinterview instrument added additional items to address NHIS content-related questions to the Reinterview instrument. Previously, the QC Reinterview had only asked respondents about whether the FR followed certain survey protocols, but not about the questions asked.

The QC Reinterview problem indicators that we used for this analysis are included in Table 2 (the remaining indicators can be seen in Appendix B). We included the reinterviewer's judgment that falsification was present as a QC Reinterview problem, and included that information in our "QC Reinterview problem flag" because interviewers can be suspected of falsification without having any discrepancy flags, or might have discrepancy flags that we did not use in this analysis.

We took as a "problem indicator" only those PANDA and Reinterview indicators that seemed to warrant particular concern, and on which we a think a manager would want to take action quickly. Our goal was to focus on problems identified by either system that would be likely to indicate poor data quality, rather than simple procedural problems.

Table 2: Indicators Monitored in QC Reinterview and Used as "Reinterview Problem Flag"

Reinterview Discrepancy
Was a laptop used for the interview?
Was the household ever contacted?
Were specific questions asked (respondents asked about questions by topic)?
Does the supervisor conducting the Reinterview suspect falsification?

2. Data Source

We combined QC Reinterview and PANDA household-level data from the 2009 and 2010 NHIS data collection years. We combined these data files into a single file with sample cases as the rows. We refer to this as our "case-level" file. Two binary variables indicated whether the sample case was identified by Reinterview, or by PANDA (1= identified, 0= not identified). These two indicator variables were summarized by a four-level variable indicating whether the interviewer was flagged in (1) QC Reinterview but not PANDA (2) PANDA but not QC Reinterview (3) both (4) or neither. We created a similar data file at the interviewer level, with each row in the data file being an interviewer.

Because we wanted to compare the two systems, we required a data file that contained cases that had the potential to be identified in either system. All interviews and interviewers are subject to being identified by PANDA, because PANDA is based on the paradata they generate during the interview process. That is, all interviewers and cases have PANDA data. However, since random and supplemental Reinterview covers only a small percentage of interviewers in any given interview period, not all interviewers or cases have Reinterview data. Thus, the data file was restricted to interviewers who had been subject to Reinterview at some point in 2009-10.

Predictor variables used in our analysis from case-management software, CAPI audit trails, survey response data, and interviewer employment records.

Our case-level analytic file includes 3873 cases for 2009. Complete and partial interviews make up 2292 of these cases, and non-interviews make up 1581. The respective numbers for 2010 are 3562 total cases, 2096 complete and partial interviews, and 1466 non-interviews. Our interviewer-level analytic data set included 393 interviewers for 2009 and 413 interviewers for 2010.

3. Methodology

Analyses proceeded in two steps. First we cross-tabulated the PANDA and QC Reinterview problem indicators in the case-level and interview-level data files. Then we built regression models predicting identification in each of the systems at each level of data separately. All results are presented in the next section.

The goal of our regression analyses was to see if any of the available interviewer, case, and respondent characteristic data we had available could predict being identified by a PANDA flag or a Reinterview flag. For each year of data separately, we ran logistic regression models predicting PANDA flags and Reinterview flags independently. We also conducted multinomial logistic regressions predicting whether a case, or interviewer in the interviewer-level models, would be identified by both PANDA and Reinterview, by one or the other, or by neither.

From the variables available, we selected those that we hypothesized would predict whether an interviewer or a sample unit would be identified as a problem by the PANDA system or the QC Reinterview system. We categorized these predictors into three theoretical causes of quality problems 1) characteristics of respondents, 2) characteristics of interviewers, and 3) characteristics of sample units or households. Respondent characteristics included whether the respondent was White, Hispanic, Male, and 65 years old or older. The only interviewer characteristic we could obtain was their experience, defined as years working as an interviewer for the Census Bureau. Sample unit characteristics included whether the case was located in an urban area, final case disposition (e.g., Interview, Refusal, Vacant), the length of time the case was in the field before a final outcome was assigned, and whether the interview had at least one section completed by phone.

In order to maximize the use of information we had in the cases and variables in our data files, we used three different data sets (i.e., case bases) in our regression models. One case base included all cases (interviews and non-interviews), but excluded respondent data that were obtained in the interview. Another included only interviews and partial interviews, allowing us to use respondent predictors. A third case base including only non-interviews was used for comparison. We were interested in whether different data sets produced different relationships with PANDA and interview flags.

Predictors used with each case base include:

All cases (full and partial interviews, and non-interviews): Length of time in the field before completion; Case dispositions, Interviewer experience; Urban case

Interviews (full and partial) only: Length of time in the field before completion; Interviewer experience; Hispanic respondent; White respondent; Older respondent

(65 years or older); Male respondent; Urban case; At least one section completed by telephone

Non-interviews only: Length of time in field before completion; Case disposition; Interviewer experience; Urban case

For our interviewer-level models we summed case-level binary predictor variables into interviewer-level indicators of the percent of an interviewer's workload with each attribute. We also created a binary variable indicating whether the interviewer had *any* cases with that characteristic. For example, if the original household-level variable was "respondent is Hispanic," we created interviewer-level variables indicating "percent of cases of Hispanic origin in the interviewer's workload," and "interviewer has any Hispanic cases in their workload (yes/no)" We developed binary indicators for respondent race, Hispanic ethnicity, age (65 or older), sex, urbanicity, and telephone interview mode. Interviewer experience was defined as having more than one year of experience at the beginning of the year. We also experimented with different recodings of the categorical predictors. We recoded continuous predictors to categorical in two additional ways. One approach was based on the median of the original distribution. For example, the median of "proportion of interviewer workload in which respondent is White" was 75%, so a binary variable was created such that a 1 was assigned if the interviewer's caseload was 75% or more, and a 0 if it was less than 75%. We also recoded continuous variables with heaping at 0% and 100% into three-category variables (e.g., 1=0%, 2=1-99%, and 3=100%). All predictor types were used in our model building.

3. Results

3.1 Crosstabulation of Problem Flags

Table 3 is a cross-tabulation of the two variables indicating whether the case was flagged in PANDA or flagged in Reinterview. This cross-tabulation is not significant at the alpha=.1 level (chisq=2.21, p=0.14). A significant relationship in this table would suggest that being flagged in one system predicts being flagged in the other. The cross tabulations of 2009 and 2010 data are similar, but we present 2010 data below.

Table 3: Cross-tabulation of Case-level PANDA Flags and QC Reinterview Flags

		PANDA		
		Not Flagged	Flagged	Row Total
Reinterview	Not Flagged	2819	641	3459
		79%	18%	97%
	Flagged	78	25	103
		2%	1%	3%
	Column Total	2896	666	3562
		19%	19%	100%

Table 4 is a cross tabulation of interviewer-level PANDA and QC Reinterview flags. We see a similar story to the case-level cross-tabulation of flags. Visual inspection of this table reveals little overlap in the cases flagged by each, and this is confirmed by the Chi-squared test of independence (p=0.29).

Aside from statistical significance, notice that in table 4 a larger proportion of interviewers were flagged by Reinterview and not PANDA (15% of cases) than PANDA and not QC Reinterview (7%). The pattern is a reversal from the case-level analysis, where more cases were flagged by PANDA and not QC Reinterview (18%) than QC Reinterview and not PANDA (2%). It is unclear if this pattern reflects any real difference or is simply an artifact of the level of summarization in each table (cases v. interviewers). The results may not be directly comparable without knowing more about the distributions of cases across interviewer workloads.

Table 4: Cross-tabulation of Interviewer-level PANDA Flags and QC Reinterview Flags

		PANDA		
		Not Flagged	Flagged	Row Total
Reinterview	Not Flagged	312	29	341
		76%	7%	83%
	Flagged	63	9	72
		15%	2%	17%
	Column Total	375	38	413
		90%	9%	100%

3.2 Regression Analysis

We ran all of our models as forced models with all predictors included and with forward inclusion, backward deletion, and stepwise selection methods. We

evaluated fit with pseudo-R^2 statistics and Hosmer-Lemeshow goodness-of-fit tests. Selection methods provided very little additional insight into the predictive ability of our variables. The models in which we replaced some of our continuous predictors with categorical predictors had better model fit than those with all continuous predictors, but probably as a function of the recoding itself rather than any true relationship between the predictors and the outcome. Our regression exploration resulted in models with overall lack of fit, even under the various scenarios described above. This was true at both the case level and interviewer level. At the case-level, maximum rescaled R^2 ranged from less than .01 to .06. The model with the highest R^2 also had a significant Hosmer-Lemeshow statistics, indicating lack of fit.

Lack of fit was seen in the interviewer-level models, as well. Our best-fitting model was a multinomial regression, with a final maximum rescaled R^2 of .117. As stated above, we think this was due to the binary recoding of predictors.

Due to the overall lack of fit across analyses, we do not present any coefficients.

4. Limitations

We were limited from the start by our categorization of PANDA and Reinterview problems. Our PANDA problems were very different from our Reinterview problems, and thus, based on this categorization of "problem cases" in each system, we would not expect much overlap in the cases or interviewers captured by the two systems. This is not an essential limitation of the data, but of the way we coded our problem flags. A true test of the research question of whether the systems capture the same or different cases would require a different coding of problem flags that are more similar across systems. Any future research on these data will modify the coding of the outcome.

Further, we only had access to a very limited set of predictor variables. If we aim to explore the role of respondent, interviewer, and case predictors in identification of problem cases, we should obtain more predictor variables to represent our potential causes of problems. Some of those predictors might come from our contact history instrument or other paradata sources.

The relationship between the two systems, and additional supervisor decision processes that can identify problems play an unmeasured role in quality assurance processes. Interviewers can be put into supplemental reinterview by selection of a supervisor who suspects that they are having a problem with their caseload. We make no attempt to model this decision process. Second, interviewers can be put into QC reinterview based on PANDA results. Statistically speaking, this would induce correlation between the cases and FR's identified as problematic by each system. If we had found a statistical relationship between the two systems, we may be concerned that it was due to this correlation. Since we found no relationship between the two systems, we can be confident that if there is a

correlation induced by FRs being identified in PANDA and sent to reinterview, it is so small that it does not influence the overall results.

Finally, a major component of the system producing these data is the interviewer, but we did not model that explicitly. We created models from a case-level file, and an interviewer-level file, but we did not model interviewer effects. This could be done by developing multi-level models with a nested data file, with cases nested within interviewers. This would be a more statistically compelling way to treat these data, and will be done in any future research with these data.

5. Discussion

Comparing data quality assurance systems is a complicated task. Systems are developed to achieve different goals and thus record different types of data in different ways. In our case, it was expected from the start that PANDA, which was developed to use paradata to monitor performance, and the Quality Control Reinterview, which was developed to catch falsification and interviewers who do not follow protocols, might catch different cases. Yet, if they are to be compared, it is imperative that they are compared on like measures, should those measures exist. Different systems with different goals may have few measures that capture identical concepts on which they can be compared, which makes a complicated analytic situation.

It is clear that the causal systems underlying data quality are complex, and thus modeling the capture of data quality is subsequently more complex. Further, the ultimate concern of quality assurance systems is systematic error in the resulting data and estimates calculated from those data. We envision using paradata from quality assurance systems like QC Reinterview and PANDA to model the relationship between problems in the field and error in resulting survey estimates. This could lead to discoveries about whether the cases we flag in our systems actually have a negative impact on the total survey error of our estimates. Much more research is needed to fully understand the independence and redundancy of quality control systems and their relationship to total survey error.

References

O'Reilly, Jim. 2009 "Paradata and Blaise: A Review of Recent Applications and Research" *12th International Blaise Users Conference, Latvia.*

Taylor, Beth L. 2009 "The 2006 National Health Interview Survey (NHIS) Paradata File: Overview and Applications". *FedCasic Conference, Washington, DC.*

Rowe, Christina. 2009 "Using Paradata to Monitor Data Quality*". International Field Director's and Technologies Conference. Delray Beach, FL.*

Sirkis, Robyn and Matt Jans, Forthcoming. “Using Statistical Process Control to Understand Variation in Computer-Assisted Personal Interviewing Data.” *Joint Statistical Meetings, Miami Beach, FL, August, 2011.*

U.S. Census Bureau. March, 2009. National Health Interview Survey PANDA Guide: Description of Quality Assurance (QA) Indicator Reports.

Appendix A: PANDA Indicators Not Used in the Current Analysis

Indicator	Description
Excessive Don't Know and Refused Answers	Complete (201) and Sufficient Partial (203) interviews where more than 50 percent of responses on key questions in the Family, Sample Adult, and Sample Child Sections are Don't Know or Refused.
Late Start Rate	Percentage of an interviewer's cases where the first contact attempt was recorded in the last week of the interview period.
Late Completion Rate	Percentage of an interviewer's cases completed in the last week of the interview period.
Completion Rate	Total number of complete interviews (Outcome Code 201) divided by the total number of Eligible Cases (Completes, Sufficient Partials, Insufficient Partials, Refusals, Language Problem, No one home, Temporarily Absent, Other) per Interviewer
Partial Rate	Total number of "sufficient partial interview" cases (Outcome Code 203) divided by the total number of Eligible Cases for each interviewer
Sample Adult No Consent to Record Linkage Rate	Percentage of cases where the Sample Adult Social Security Number (SSN) and record linkage consent items were asked and were answered with Don't Know or Refused
Breakoff Rate	Total number of Breakoff Cases (where the interview was stopped in progress) divided by the total number of cases with Outcome Codes of 201 (Complete), 203 (Sufficient Partial) and 215 (Insufficient Partial)
No Telephone Number/ Alias Used in 201/203/236 Cases	Percent of 201 (Complete), 203 (Sufficient Partial), 236 (Screened-Out) cases missing a telephone number, an alias was reported, or both.
Telephone Rate	The percentage of complete (201) and partial (203) interviews where one or more main sections of the NHIS are completed by telephone.
Screened-Out Rate	The percentage of contacted cases that screened out. (Having at least one nonmilitary household member who is Black, Hispanic, or Asian
Type B/C Rate	The percentage of Type B and Type C cases out of total workload. (Percentage of Noninterview Cases)
Family Interview Overnight Start Rate	The percentage of Complete (201) cases where the Family Interview was started between 11:00 p.m. and 7:00 a.m. local time.
Sample Adult Interview Overnight Start Rate	The percentage of Complete cases where the Sample Adult Interview was started between 11:00 p.m. and 7:00 a.m.
Missing CHI	The percent of all cases worked by the interviewer that did not have at least one Contact History Instrument record.

Appendix B: Reinterview Indicators Not Used in the Current Analysis

Reinterview Discrepancy
Original interview status incorrect
The Type B/C status by observation is incorrect
Type A status is incorrect
The interviewer classified the interview or Type A housing unit as a Type B/C
Household roster is incorrect
The interviewer conducted telephone interview instead of a personal visit
The interviewer entered a bad telephone