



Child Care and Development Fund Using the ACF-801 Data Assessment Report

Technical Bulletin #9r-v2

REVISED: August 2010, Reviewed August 2011

I. INTRODUCTION

This Technical Bulletin describes how to use the Summary Data Assessment Report (**Summary Report**). Designated staff in States and Territories automatically receive this report via e-mail after their ACF-801 case-level data has been submitted to the Child Care Bureau (CCB) and has been processed in the Child Care Bureau Information System (CCBIS).¹ Note that the Detailed Assessment Report no longer is generated by CCBIS and grantees may instead use the *Child Care Data Viewer* software to capture similar information.

The Summary Report assists States and Territories in correcting data submissions by providing information about potential data or submission errors. This report serves as just one tool that grantees should use to ensure that complete and accurate CCDF program information is submitted to CCB. The data submitted will be reported to Congress by the Secretary of the Department of Health and Human Services, used in national child care research, and may serve as the basis for programmatic decisions at the Federal and State level. Thus, it is important that States and Territories submit accurate and complete data.

The remainder of this Bulletin is organized in the following sections:

Section II: Overview of the CCBIS Data Check Process

Section III: The Summary Report

Part A: File Format Assessment

Part B: Submission Summary

Part C: Data Quality Assessment

Part D: Cross File Checks

Section IV: The Child Care Data Viewer

Section V: Summary and Conclusions

¹ The Summary Report is sent via e-mail to the contact at the child care agency identified in the file's summary record, to other program staff the grantee identifies, and to child care contacts in the ACF Regional Offices and other officials in the Child Care Bureau.

II. OVERVIEW OF THE CCBIS DATA CHECK PROCESS

Once data are transmitted to the CCBIS, they are passed through a series of quality checks to validate the formatting and the content of the data. These checks are conducted in the following order and data must pass each check before proceeding to the next check:

1. **File Format Check** - checks the length of the records and determines whether all the necessary parts for each record are present. (See the latest version of Technical Bulletin #4 for the file standards used.) The actual content of the data is not checked at this stage.
2. **Data Quality Check** - checks the content of properly formatted records. (See the latest version of Technical Bulletin #3 for further discussion of the data standards used.) If at any stage data fail the check, then the remaining sequence of data quality checks are not conducted on the data (for instance, if, in a properly formatted record, the data for an element are missing, then the out-of-range, internal consistency, and cross file checks are not conducted for that data):
 - **Missing** - checks the data in properly formatted files to ensure that there are values present (rather than blanks) for the required data elements.
 - **Out-of-Range** - checks to ensure that the data which passed the file format and missing checks are within the range of valid values.
 - **Internal Inconsistency** - checks to ensure that values of data elements which passed the file format, missing, and out-of-range checks are consistent with other related data elements.
3. **Cross File Check** - checks all the data in the entire submission and determines whether the characteristics of the data submission conform to standards that are *generally* true of child care programs, though they may not be true for an individual State.

III. THE SUMMARY REPORT

The Summary Report² presents the results of the CCBIS Data Checks described in Section II. The results presented here can help you find and correct errors in your submission. These problems may be the result of data entry or data extraction errors.

```

CCBIS8018EM
Date: Wednesday, 06/08/2010 03:10PM
File Name: VGLIWI.PCCIS.NDM.YS.HUB.Y2005.D0608.T1000
State: YourState

                ACF-801 Data Submission
                Summary Data Assessment Report
                Report Date: APRIL 2010

                File Format Assessment

2 Families were rejected because of format errors
2 Children were rejected because of format errors
2 Settings were rejected because of format errors

3 Families were submitted without any children
3 Children were submitted without any settings
3 Family records with blank case identifiers and SSNs

                Submission Summary

Families receiving subsidized child care (per summary record):    3,036

Families:                3,036
Children:                5,053
Settings:                5,148
Footnotes:                1

                Data Quality Assessment

Data Element      Missing  Out-of-  Internally  Total  Subject  Success
                  Data      Range   Inconsistent  Errors Records  Rate
Family Data:
01: Rpt Period    0        0        0          0     3,036  100.00%
02: State ID      0        0        0          0     3,036  100.00%
03: SSN           0        0        0          0     3,036  100.00%
04: FIPS Code     0        0        0          0     3,036  100.00%
05: Single Parent 0        0        0          0     3,036  100.00%
06: Care Reason   0        0        0          0     3,036  100.00%
07: Copay         0        0        235        235   3,035  92.25%
08: Subsidy Begin 0        0        0          0     3,036  100.00%
09: Income        0        0        235        235   3,035  92.25%
10: Employment   0        0        0          0     3,035  100.00%
11: TANF         0        0        0          0     3,035  100.00%
12: TANF MOE     0        0        0          0     3,035  100.00%
13: Housing/Cash 90       0        0          90    3,035  97.02%
14: Food Stamps  0        0        0          0     3,035  100.00%
15: Other Assist 0        0        0          0     3,035  100.00%
16: Family Size  0        1        77         78    3,036  97.43%
    
```

Part A: File Format Assessment - lists the number of family, child, and setting records which were not formatted correctly.

Part B: Submission Summary - lists the number of family, child, setting, and footnote records that passed the file format checks and were submitted for the data quality checks.

```

Child Data:
17: Child SSN     0        0        0          0     5,053  100.00%
18: Ethnicity    0        0        0          0     5,053  100.00%
19: Amer. Indian 0        0        0          0     5,053  100.00%
20: Asian        0        0        0          0     5,053  100.00%
21: African Amer. 0        0        0          0     5,053  100.00%
22: Pac. Islander 0        0        0          0     5,053  100.00%
23: White        0        0        0          0     5,053  100.00%
24: Gender       0        0        0          0     5,053  100.00%
25: Birth Date   0        0        0          0     5,053  100.00%

Setting Data:
26: Care Type    0        0        0          0     5,148  100.00%
27: Amount Paid  0        0        0          0     5,148  100.00%
28: Hours of Care 0        0        0          0     5,148  100.00%

-----
Totals:          90        1        547        638
    
```

Part C: Data Quality Assessment - lists the results of the checks on each data element for missing, out-of-range, and inconsistent data.

```

                Cross File Checks

Please note:
-----
Element 06: No records indicate Option 3: "Both Employment and Training/Education
as the reason for receiving subsidized child care.
Element 04: No data submitted for the following counties:

County Name      FIPS Code
-----
RAVENSIDE        24002
KING GEORGE      24051

Element 11: All records have the same value
Element 11: No records indicate "TANF Income"
Element 14: All records have the same value
Element 14: No records indicate "Assistance under the Food Stamp Act of 1997"
Element 15: All records have the same value
Element 15: No records indicate "Other Income Sources"
Element 18: All Records have the same value
Element 18: No records indicate "Hispanic/Latino" ethnicity
Element 19: Only 0.27% of all records indicate "American Indian or Alaskan Native"
race
Element 20: Only 2.65% of all records indicate "Asian" race.
    
```

Part D: Cross File Checks – lists possible errors in the submission from a content and programmatic standpoint.

² To ensure that the report is easily readable, e-mail font should be set to courier 10 point font with left and right margins of one inch or less.

Part A: File Format Assessment

The File Format Assessment in the Summary Report presents the results of the formatting check. An example of a File Format Assessment is displayed in Figure A, below. Here, each record is checked against the file format standards, as found in the latest version of Technical Bulletin #4.

File Format Assessment	
2	families were rejected because of format errors
2	children were rejected because of format errors
2	settings were rejected because of format errors
3	families were submitted without any children
3	children were submitted without any settings
3	family records with blank case identifiers and SSNs

Figure A: File Format Assessment

Note that there are other file format errors which disable the CCBIS Data Checks. ACF-801 data will not be processed if any of the following occur:

- The summary record is omitted
- Monthly data within quarterly submissions are submitted out of chronological order (e.g., May, April, June)
- Some combination of monthly files other than one single month, or three months in a defined quarter (e.g. January, February, and March) are submitted at one time.

See Table 1 on the following page for a list of the messages that may appear in the File Format Assessment, what each message means, and possible ways to address issues brought up by each.

Error message	File Format Standard	Issues to Address
<i># families were rejected because of format errors</i>	Each family record should begin with an "F" delimiter, followed by 60 record positions with family data.	Check the record for the following: - The record length is either too short or too long due to missing data, incorrectly placed data or extra values. - The "F" delimiter may be missing. If all the records in the submission are incorrectly formatted, this may indicate a problem with mapping or extraction of the data.
<i># children were rejected because of format errors</i>	Each child record should begin with a "C" delimiter. The first child record should start at the 62nd position in the record, followed by 22 record positions with child data.	Check for an error with respect to record length or the delimiter (see above).
<i># settings were rejected because of format errors.</i>	Each setting record should begin with an "S" delimiter, followed by 9 record positions with setting data.	Check for an error with respect to record length or the delimiter (see above).
<i># families were submitted without any children</i>	Each record (which represents a household receiving subsidized child care) must contain one family record and at least one child record.	Check the logic of your extraction routine to see why no children were associated with the family.
<i># children were submitted without any settings</i>	Each child record must have at least one setting record.	Check the logic of your extraction routine to see why there are no child care settings recorded for the child.
<i># family records with blank case identifiers and SSNs</i>	Each family record must contain an SSN or a unique state identifier. When both are missing, it will not be possible to distinguish this record from the others, so no data quality check will be performed on this data.	Check your extraction routine to insure it is capturing SSNs or case identifiers for each family. Also, check that every family in your system has an SSN and/or unique state identifier.

Table 1: File Format Assessment Messages

Part B: Submission Summary

The Submission Summary contains information from two different parts of the ACF-801 file. In Figure B, below, notice the first line is:

```
Families receiving subsidized child care (per summary record): 3036
```

As indicated, the total number of families served during the month is taken from the header (or summary) record that precedes the State’s case-level submission. Other information in the Submission Summary includes counts of records accepted for processing by the Child Care Bureau Information System (CCBIS) after the file format checks are run. Any family, child, or setting record that fails the file format checks is not counted. Only correctly formatted records will be checked for data quality, the results of which appear in the next section, Data Quality Assessment.

NOTE: If a State submits a full population, **and** if all family records are properly formatted, the number of families reported in the header record should equal the count of families processed by the CCBIS. If a State submits sample data, the two numbers will differ, as in the example below.

Submission Summary	
Families receiving subsidized child care (per summary record):	3036
Families:	204
Children:	350
Settings:	398
Footnotes:	2

Figure B: Submission Summary

See Table 2 on the following page for a description of counts that appear in the Submission Summary, what each message means, and possible ways to address potential problems.

Counts	Definition	Potential Problems
<i>Families receiving subsidized child care (per summary record):</i>	<p>The total number of families served as reported on the summary record. For example:</p> <p>"M20100400030360000334#Jane Smith#(703) 555-1234 ext.666#(703) 555-876#jsmith@dhr.anystate.us."</p>	<p>If this number does not equal the total number of families that received subsidized child care as reported in the case-level data during the reported month, correct the number and re-submit the data.</p>
<i>Families:</i>	<p>The number of properly formatted family records counted by the system, which are then submitted for the data quality checks (any invalid/improperly formatted family records are not included in this count).</p>	<p>If the number of family records indicated here is inconsistent with the number of families listed in the submission summary (unless your State submitted sample data, in which case there should be a minimum of 200 family records), determine where the error is, make corrections, and resubmit the file.</p>
<i>Children:</i>	<p>The number of properly formatted child records that were counted by the system and submitted for data quality checks.</p>	<p>The number of child records processed by CCBIS and checked for format and other errors should match the number of submitted child records. If there are fewer records, identify the error(s), make corrections, and resubmit the file.</p>
<i>Settings:</i>	<p>The number of properly formatted setting records that were submitted for data quality checks.</p>	<p>There should be at least one setting record for each child. If the number of settings does not equal or exceed the number of children served, identify the error(s), make corrections, and resubmit the file.</p>
<i>Footnotes:</i>	<p>The footnotes that were counted. These too must be correctly formatted, beginning with a number indicating the data element to which it refers. (See the latest version of Technical Bulletin #4.)</p>	<p>If this number does not equal the number of footnotes your State intended to submit, examine the footnotes to locate the problem, correct it, and resubmit the file.</p>

Table 2: Description of Counts on the Submission Summary

Part C: Data Quality Assessment

The Data Quality Assessment contains the results of data quality checks for missing, out-of-range, or internally inconsistent values. A sample of a section of the Data Quality Assessment is displayed in Figure C, below.

Data Quality Assessment						
Data Element	Missing Data	Out-of-Range	Internally Inconsistent	Total Errors	Subject Records	Success Rate
Family Data:						
01: Rpt Period	0	3	0	0	13,034	99.98%
02: State ID	0	0	0	0	13,034	100.00%
03: SSN	0	0	0	0	13,034	100.00%
04: FIPS Code	0	0	0	0	13,034	100.00%
05: Single Parent	0	0	0	0	13,034	100.00%
06: Care Reason	0	0	0	0	13,034	100.00%
07: Co-pay	0	0	256	256	13,027	98.03%
08: Subsidy Begin	0	0	0	0	13,034	100.00%
09: Income	0	0	256	256	13,027	98.03%
10: Employment	0	0	0	0	13,027	100.00%
11: TANF	0	0	0	0	13,027	100.00%
12: TANF MOE	0	0	0	0	13,027	100.00%
13: Housing/Cash	908	0	0	0	13,027	93.03%
14: Food Stamps	0	0	0	0	13,027	100.00%
15: Other Assist	0	0	0	0	13,027	100.00%
16: Family Size	0	0	501	501	13,034	96.15%

Figure C: Data Quality Assessment

The columns listed in this section are as follows:

- Data Element:** The 28 data elements for the ACF-801 are listed in the first column. They are grouped into three sections beginning with the “Family Data” section (elements #1-16), “Child Data” (elements #17-25), and “Setting Data” (elements #26-28).
- Missing Data:** The number of records containing no values (blanks) for a data element.
- Out-of-Range:** The number of records where values for the data element were present but outside the allowable range of values according to the CCBIS data standards (See the latest version of Technical Bulletin #3).
- Internally Inconsistent:** The number of records where values for a data element were entered and in the correct range of values, but conflict with the values for another, related data element (See the latest version of Technical Bulletin #3).

- **Total Errors:** The total number of missing, out-of-range, and internally inconsistent errors found for each element.
- **Subject Records:** The number of records that were tested for data quality errors. The number of family, child, and setting records should be the same as those listed in the submission summary. (Note that the only exception is protective services cases, which do not undergo the missing, out-of-range, and internal inconsistency checks for those elements relating to co-payment (element #7) or income (elements #9-15).
- **Success Rate:** The percentage of records that passed the missing, out-of-range, and consistency checks (“Subject Records” minus “Total Errors”, divided by “Subject Records”).

NOTE – The Summary Assessment process checks only the information that has been submitted in the ACF-801 report. It is up to each grantee to assure that data contained within that report accurately represent its own CCDF program. Large numbers of missing, out-of-range, or inconsistency errors reported in this section *may* indicate a systematic problem with the preparation of the submission. For instance:

- If more than 5% (success rate $\leq 95\%$) of the records have a particular element missing, this might indicate that: (1) there is an error in the software program used to pull the required information from another database and reformat it to meet the Federal standard, or (2) the database does not contain the data and/or that the data is not being collected. Check with your programmers to see what may be the exact cause.
- If more than 5% (success rate $\leq 95\%$) of the records have a particular data element out-of-range, this could indicate a possible problem with mapping or extraction of the data. For example, the extraction program may inadvertently transform the source’s database value of “F” (female) to “4” instead of the required Federal value of “2.” Check with your programmers to see what may be the exact cause.
- If more than 5% (success rate $\leq 95\%$) of the records have a particular data element noted as internally inconsistent, this could indicate a possible problem with data entry, or any of the above mentioned software program issues. For example, if an applicant is reported as a single parent, the family size must be ≥ 2 (one parent with at least one child). A family size of one (1) would trigger an inconsistent error message. Check with your data entry staff and programmers to see what may be the exact cause.

Part D: Cross File Checks

The Cross File Checks (see Figure D, below) present the result of a programmatic analysis of all the data that passed the file format and data quality checks. While data for individual records may be correct, at times the data for a submission as a whole may not conform to general programmatic norms. When the values for a particular data element fail to meet a general programmatic expectation, this triggers a “red flag” and a notice is produced for that cross-file check. Check these notices to see if they apply to your program. Table 3 on the following page contains a list of all the notices that may appear (not all cross-file checks are conducted on sample submissions; these are indicated by an asterisk).

NOTE: These standards are based on *general* programmatic characteristics for State child care programs. Not all may be applicable to your specific State.

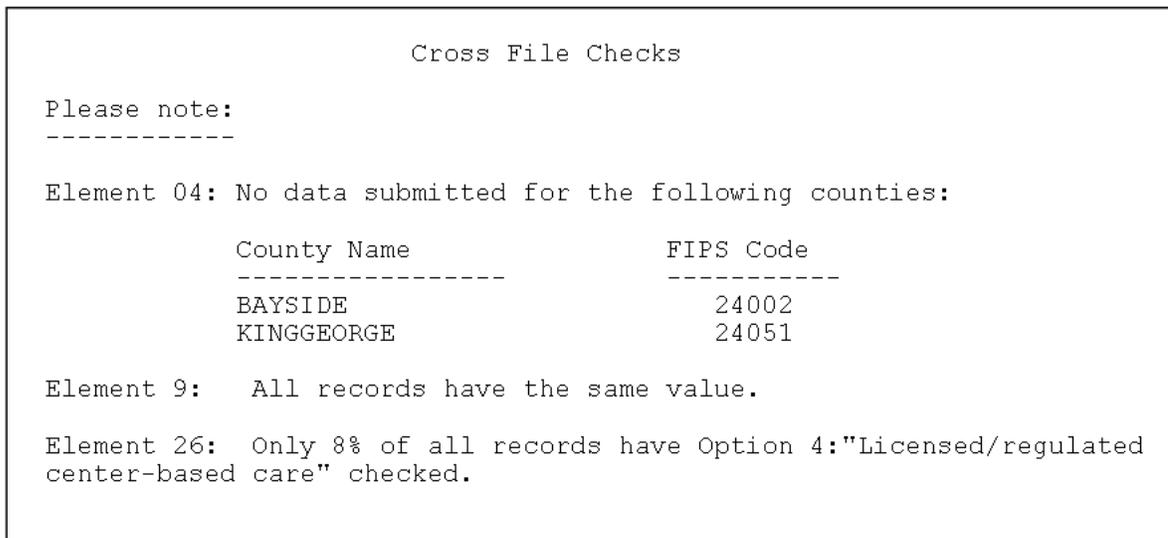


Figure D: Cross File Checks

In the above full population example, the following programmatic checks were triggered:

- For Element #4, FIPS Code, counties which did not submit case records are listed – this may indicate missing information.
- For Element #9, Employment Income, all the records contained the same value for this element– it is unlikely that all families would have exactly the same income – this may indicate an extraction or other system error.
- For Element #26, Type of Child Care, 8% of records indicated licensed/regulating center-based care as the type of child care provider. For most States, this percentage is considerably higher, so you would want to confirm that this figure is true for your State.

Cross File Check Messages - Conducted on Full Record

Your State submitted no families with more than one child receiving services.
 Your State submitted no families with more than two children receiving services.
 Your State submitted no children receiving child care services in more than one setting for the month.
 Your State submitted no children receiving services in more than two settings for the month.

Data Element	Cross File Check Messages - Conducted on Data Elements Asterisks (*) indicate checks <u>not</u> conducted on sample submissions
01: Report Period	- Month/Year of Report Period is not the same as that of the Summary record.
02: Unique State Identifier	- No identifiers submitted. Please consider submitting optional case identifiers.
03: SSN	- No checks.
04: FIPS Code	- No data submitted for the following counties: Name/FIPS Code.* - x % of all records have the same FIPS code, County Name.*
05: Single Parent	- None of the records indicate Single Parent Status. - Only x% of all records indicate Single Parent status.
06: Reason for Care	- None of the records indicate Option 1: "Employment, including on-the-job training" as the reason for receiving subsidized child care. - None of the records indicate Option 2: "Training/Education" as the reason for receiving subsidized child care. - None of the records indicate Option 3: "Both Employment and Training/Education" as the reason for receiving subsidized child care. - x % of all records indicate Option 4: "Protective Services" as the reason for receiving subsidized child care. - All records have the same value.
07: Family Co-pay	- All records have the same value.
08: Subsidy Began	- All records have the same value. - Element 01 and Element 08 have the same value in x% of the records.
09: Family Income	- All records have the same value.
10: Employment	- All records have the same value.
11: TANF	- All records have the same value. - None of the records indicate "TANF Income." - Only x% of all records indicate "TANF Income."
12: State Dollars counted towards TANF MOE	- No Checks.
13: Housing Voucher or Cash Assistance	- All records have the same value. - None of the records indicate "Housing Voucher or Cash Assistance." - Only x% of all records indicate "Housing Voucher or Cash Assistance."
14: Assistance Under the Food Stamp Act of 1977	- All records have the same value. - None of the records indicate "Assistance under the Food Stamp Act of 1977." - Only x% of all records indicate "Assistance under the Food Stamp Act of 1977."
15: Other Federal Cash Income Programs (e.g. SSI)	- All records have the same value. - None of the records indicate "Other Income Sources." - x% of all records indicate "Other Income Sources."
16: Family Size	- All records have the same value.
17: Child SSN	- No Checks.

Data Element	Cross File Check Messages - Conducted on Data Elements Asterisks (*) indicate checks <u>not</u> conducted on sample submissions
18: Hispanic/Latino	<ul style="list-style-type: none"> - All records have the same value.* - None of the records indicate "Hispanic/Latino" ethnicity. - Only x% of all records indicate "Hispanic/Latino" ethnicity.
19: Am.Ind./AK Native	<ul style="list-style-type: none"> - All records have the same value.*
20: Asian	<ul style="list-style-type: none"> - All records have the same value.* - None of the records indicate "Asian" race. - Only x% of all records indicate "Asian" race.
21: Black/African-Am.	<ul style="list-style-type: none"> - All records have the same value.* -None of the records indicate "Black or African American" race. - Only x% of all records indicate "Black or African American" race.
22: Hawaiian/Pacific Isl.	<ul style="list-style-type: none"> - No checks.
23: White	<ul style="list-style-type: none"> - All records have the same value. - None of the records indicate "White" race.* - Only x% of all records indicate "White" race.
24: Gender	<ul style="list-style-type: none"> - None of the records indicate Option (2): "Female" as child gender. - x% of all records indicate Option 2: "Female" as child gender.
25: Birth Mo./Yr.	<ul style="list-style-type: none"> All records have the same value.
26: Type of Care	<ul style="list-style-type: none"> - All records have the same value. - Less than 10% of all records have Option 1: "Licensed/regulated in-home child care" checked. - Less than 10% of all records have Option 2: "Licensed/regulated family child care" checked. - Less than 10% of all records have Option 3: "Licensed/regulated group home child care" checked. - None of the records have Option 4: "Licensed/regulated center-based care" checked. - Only x% of all records have Option 4: "Licensed/regulated center-based care" checked.
27: Total Payment	<ul style="list-style-type: none"> All records have the same value.
28: Hours of Care	<ul style="list-style-type: none"> All records have the same value.

Table 3: Cross File Check Messages (cont.)

IV. THE CHILD CARE DATA VIEWER

The Summary Report also can be generated using the *Child Care Data Viewer (Data Viewer)*. The *Data Viewer* software was developed by the Child Care Bureau (CCB) to give State and Territory CCDF grantees the ability to assess the quality of data files prior to transmitting them to the CCB. It is Microsoft Access based computer software that is used in conjunction with prepared ACF-801 monthly data files. The *Data Viewer* serves two primary purposes: 1) checking the quality of data, and 2) generating reports.

Data Viewer functions include:

- The generation of the Summary Report which provides a high level quality assurance (QA) check of your monthly data,
- The ability to **view specific case records** for a more detailed review,
- Display of printable family record information to facilitate corrections, and
- Automatic calculation of standard reports (based on a grantee's monthly data) that can be used by program staff for supervision, public inquiries, resource development, program improvement, and community education.

The *Child Care Data Viewer Users' Manual* is available online at:

<http://www.acf.hhs.gov/programs/ccb/ta/ccarc/guide/ccdv.doc>

<http://www.acf.hhs.gov/programs/ccb/ta/ccarc/ccdv.htm>

V. SUMMARY AND CONCLUSIONS

This Bulletin has presented information on how to use the Summary Report that is returned to grantees after their ACF-801 report is submitted and processed. The Summary Report provides a view of the grantee data as a whole, and allows States/Territories to assess potential errors with extraction, mapping, or data collection and entry. Since this information will be used in research, reported to Congress, and may serve as the basis for programmatic decisions on child care, it is important that submitted data are accurate and complete.

For more information about the data quality and file format standards used by CCBIS for ACF-801 case-level data, see the latest versions of Technical Bulletins #3 and #4 on the CCB website (www.acf.hhs.gov/programs/ccb/report/formhelp/techbull/index.htm).

If you have questions and need additional assistance, contact the Child Care Automation Resource Center (CCARC) Monday-Friday between 9:00 am and 5:00 pm prevailing Eastern Time:

Phone (toll free): 1-877-249-9117

E-mail: ccarc@childcaredata.org