In collaboration with the Medicaid Program Office, 
Agency for Health Care Administration, and 
the Department of Health

Explanation of Apparent Discrepancies Among 
Different Editions of Maternal Infant Reports

Amended October 25, 2002
Explanation of Apparent Discrepancies
Among Different Editions of
Maternal Infant Reports

Produced by:
Maternal Child Health and Education Research Data Center
at the University of Florida
a Branch of the Lawton and Rhea Chiles Center
for Healthy Mothers and Babies
College of Public Health
University of South Florida

In conjunction with

The Florida Agency for Health Care Administration
Rhonda M. Medows, MD, FAAFP, Secretary

and

The Florida Department of Health
John O. Agwunobi, M.D., M.B.A., Secretary
Table of Contents

Explanation of Apparent Discrepancies Among Different Editions of Maternal Infant Reports

Table of Contents .......................................................................................................................... iii
Acknowledgements ...................................................................................................................... iv
Background .................................................................................................................................. 1
Frequently Asked Questions ......................................................................................................... 4
Methodology ................................................................................................................................. 6
Quality Assurance Issues ............................................................................................................. 9
Acknowledgements

Project Director

Michael B. Resnick, Ed.D.

Maternal Child Health and Education Research and Data Center, a branch of the Lawton and Rhea Chiles Center for Healthy Mothers and Babies, would like to acknowledge the following individuals for their diligent work and contributions to this report:

**University of Florida Faculty and Staff**

Mario Ariet, Ph.D.  
Ananthakrishna Babu, B.S.  
Randy L. Carter, Ph.D.  
Hao-yi Chen, Ph.D.  
Rajeeb Das, M.S.P.H.  
David Figlio, Ph.D.  
Doug Grove, M.S.  
Christina Hayes, B.A.  
Changxing Ma, Ph.D.  
Lisa Newman, B.A.  
Steven Priest, B. S.  
Jeffrey Roth, Ph.D.  
Suzanne J. Scott, Ph.D.  
Anuradha Vesudevan, B.S.  
Sam Wu, Ph.D.

**University of South Florida Faculty and Staff**

David Darr, M.D.  
Edward A. Feaver, M.Div., M.S.E.  
Michael Graven, M.D.  
Charles S. Mahan, M.D.

We would also like to thank the following individuals for contributing their support and expertise to this report:

**Florida Department of Health**

C. Meade Grigg, M.A.  
Karen Freeman, B.A.  
Jane Menges, M.A.  
Rhonda White, M.B.A.

**Agency for Health Care Administration**

Susan Chen, B.S.  
Nancy L. Ross, Ph.D.
Background

This brief report summarizes the principal changes that have been introduced into the two annual maternal/infant reports that are produced for AHCA by the Maternal Child Health and Education Research and Data Center (MCHERDC) at the University of Florida (UF). These reports began to be issued by MCHERDC under a sub contract with the Chiles Center starting in 2000. Prior to 2000, these annual reports were produced at the University of South Florida under the direction of Patricia A. Gorzka, PhD.

In 2000, the first year UF was contracted to produce the reports, no changes were made to the computer programs that generated the numbers in the tables that accompany each pregnancy outcome indicator. In a first step toward understanding how these computer programs had been constructed and how their output was generated, MCHERDC contracted with a USF programmer (Nancy Lemrow, PhD) to help us understand the working of USF’s complex merging program that creates the analyzable data set used to produce the maternal-infant reports. The analyzable data set is termed “File28” because it was originally produced as the end product of a 28-step sequence of merges. Early in the period of transition from USF to UF, MCHERDC programmers (Yuanshan Sun, PhD and Darryl Marsee, MS) noted inconsistencies and inefficiencies in the inherited program code, but because of time constraints no changes were made to the algorithms prior to production of the reports issued February 1, 2001 that covered the 5-year period 1995-1999 (for Health Status Indicators) and the 2-year period 1998-99 (for Medicaid Managed Care).

Starting in 2001, MCHERDC programmers (Doug Grove, MS and Rajeeb Das, MPH) introduced three changes to the reports as a result of discovering and correcting errors in File 28. These changes were described in the Introduction, Limitations, and Methods sections of the two reports. Briefly, the changes involved: 1) classifying women into Cash Assistance or No Cash Assistance based not on the category they were in at time of their delivery (as USF’s program had done) but rather on the category they had stayed the longest in during their pregnancy [see attached Memo of May 14, 2001 between Susan Chen and Sam Wu, co-signed by Nancy Ross]; 2) restricting the Individual HMO table to those High Exposure women who were in one particular HMO
for 180 days or 2/3 of their pregnancy and 3) reporting a group of low incidence
providers together as a single provider. Two other format changes were introduced in the
revised version of the 1995-1999 report (cover page-dated February 1, 2001 but
distributed June 20, 2001: 1) column numbers and percent for Low Exposure women in
individual HMOs for 1998 and 1999 were deleted; and 2) AFDC and SOBRA women
were relabeled Cash Assistance and No Cash Assistance. This name change actually
involved reclassifying women. The previously cited May 14, 2001 memo from Susan
Chen included a new definition of the SOBRA (No Cash Assistance) group different
from that which had been used in previous reports. In earlier annual reports produced by
USF, Medicaid had been divided into AFDC and SOBRA, which left a considerable
number of Medicaid women in neither category. When the decision was made to drop
the AFDC and SOBRA nomenclature which had become obsolete as a result of welfare
reform legislation, thousands of women who formerly had not been assigned to either
SOBRA or AFDC were distributed to the Cash Assistance and No Cash Assistance based
on the eligibility Reports prepared for 2002 reflect further discovery and correction of
errors found in the algorithms creating File 28. The primary changes made to the 1996-
2000 Health Status Indicators report are detailed in the expanded Methodology and Data
Assurance Issues sections, which have been reproduced at the end of this report. In brief,
the File 28 code as inherited from USF, contained a programming error that counted as
Medicaid eligible women who may or may not have been Medicaid eligible during their
pregnancy. Thus, for the purposes of a birth outcomes report, it was necessary to define
Medicaid eligibility correctly as occurring only during the window of pregnancy, that is,
for 280 days back from delivery date. Setting a closing date for women whose eligibility
date had formerly been left open revealed another imprecision in File 28: an overlap in
pregnancy years. Women were being double counted if their pregnancy crossed two
calendar years. This error was corrected by truncating eligibility at the end of each
calendar year.

In the 1999-2000 Medicaid Managed Care report (revised April 8, 2002), two
minor changes were made: the High Exposure column in the third, Individual HMO
table, comprise a subset of women who were included in the second table, Medicaid
Programs, High Exposure, that is, they had to meet the criterion of being in an HMO for 180 days. Given that constraint, women in Table 3 also had to be an individual HMO for either 180 days or two-thirds of their pregnancy (in order to capture women who delivered prematurely). HMOs with less than 30 deliveries for a particular outcome indicator were not listed.

In the following Frequently Asked Questions section, several concerns which have been raised by the Medicaid office during the report review process are reproduced and answered. It is hoped that these answers will serve as the basis for further discussions about how best to present historical data that has been subject to correction and refinement.

Because it is customary for MCHERDC to issue draft versions of its reports for review by the Medicaid program office, it is imperative that these draft versions be discarded when subsequent final versions are produced. In years past, unfortunately, the distinction between draft and final versions was not clearly indicated on the reports’ title pages. In an effort to remedy this potential for misunderstanding, final versions of last year and this year’s reports have been re-issued in bound copies. All previous print and electronic versions of these annual reports should be discarded.
Frequently Asked Questions


A: An error in File 28 allowed women to be counted as Medicaid eligible regardless of whether they were actually Medicaid eligible during their pregnancy (that is, they were never closed after initial determination) was discovered in 2001. This error was noted in 2001 and corrected in 2002. The correction consisted of classifying women as Medicaid eligible only for the period of 280 days back from their delivery date. This recalculation, based on the 280-day pregnancy window, caused reassignment of some pregnant women from No Cash Assistance to the Cash Assistance category. This is the reason for the differences between this year’s report and the second version of last year’s report. However, the first version of last year’s report differed substantially from the current year’s numbers. The reason for the change between the two versions of the 1995-1999 report, as explained earlier in this report, was due to the redefining of the No Cash Assistance (SOBRA) group (as stated in the memo of May 14, 2001).

Q. Why do the numbers of deliveries to Hispanic women in the 1996-2000 Health Status Indicators Reports differ from the numbers reported in the 1995-1999 edition of the Health Status Indicators Report?

A. After producing a special ad-hoc Racial/Ethnic Comparison Report in February 2002, it was decided to align the numbers that appeared there with those that would appear in the June 2002 edition of the Maternal Infant Health Status Report. The Racial/Ethnic Comparison Report followed Birth Vital Statistics on matters of race/ethnicity and counted deliveries to Black Hispanics as part of the total deliveries to Hispanic women. Thus, the numbers for Hispanic deliveries in the 2002 Health Status Indicators Report appear larger and the number of Black deliveries appear smaller.
compared to previous editions of these annual reports where deliveries to Black Hispanics were counted as deliveries to Black women.

Q. How should information be extracted from the current Race/Ethnicity Report to produce the summary PBSquare Report?

A. The following table is drawn from data available in the current Race/Ethnicity Report. Only the column for Percent Total Deliveries required calculating the proportion of each Racial/Ethnic Group to the Total Statewide as given on page 2 of the Race/Ethnicity Report.

Racial and Ethnic Data for Selected Birth Outcomes, 2000

<table>
<thead>
<tr>
<th>Racial/Ethnic Group</th>
<th>Percent Medicaid Deliveries</th>
<th>Percent Total Deliveries</th>
<th>Percent Inadequate Prenatal Care</th>
<th>Percent Low Birth Weight</th>
<th>Percent Reported Smoking During Pregnancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>African American</td>
<td>69.2</td>
<td>20.5</td>
<td>16.4</td>
<td>11.9</td>
<td>4.9</td>
</tr>
<tr>
<td>Hispanic</td>
<td>50.7</td>
<td>21.8</td>
<td>10.0</td>
<td>6.5</td>
<td>2.4</td>
</tr>
<tr>
<td>White</td>
<td>32.8</td>
<td>54.9</td>
<td>6.5</td>
<td>5.8</td>
<td>15.3</td>
</tr>
<tr>
<td>Other/Unknown</td>
<td>25.9</td>
<td>2.8</td>
<td>7.5</td>
<td>8.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Overall</td>
<td>44.0</td>
<td>100.0</td>
<td>9.3</td>
<td>7.3</td>
<td>10.0</td>
</tr>
</tbody>
</table>

Q. Why do the Health Status Indicator Reports now report only one significant decimal?

A. In an effort to make the tables of the Health Status Indicators Report more closely resemble those produced by Birth Vital Statistics, MCHERDC adopted the format used by the Florida Department of Health as well as other state and federal agencies.
Methodology

Data Sources

Data for the two Health Status reports are derived from the following sources:

- Medicaid eligibility, HMO enrollment, and Medipass files from the Agency for Health Care Administration (AHCA)
- Birth and infant death records from the Office of Vital Statistics, Florida Department of Health
- Healthy Start Prenatal Risk and Postnatal Risk Screens from the Maternal Child Health Bureau, Florida Department of Health
- WIC (Women, Infant, and Children Supplemental Nutrition Program) eligibility files, Florida Department of Health

Exclusion Criteria

Files were obtained from Vital Statistics for Florida for births and deaths from 1996 through 2000. Records meeting any of the following conditions were excluded:

1) non-resident births (mother’s residence not Florida);
2) second and subsequent children of multiple births;
3) births with a missing identifier (mother’s social security number); and
4) duplicate birth certificates;

Missing Values

Records with missing values were excluded from the summary statistics tables.

Matching Medicaid Data to Vital Statistics

Agency for Health Care Administration (AHCA) Medicaid data were matched to the Maternal Child Health core dataset using only mother's social security number; mothers missing their social security numbers were excluded. In the future matching will be improved by including the use of name and demographic information, which will allow the matching of women missing an SSN.
Categorization of Subpopulations

The total number of annual, non-excluded deliveries was divided into two categories: “Medicaid” and “Non-Medicaid.” Women who were in Medicaid for at least one day during their pregnancy were placed into the “Medicaid” category. Those with no exposure to Medicaid during their period of pregnancy were placed in the “Non-Medicaid” category.

The “Medicaid” category was further subdivided into two groups: Cash Assistance and No Cash Assistance. The No Cash Assistance group (formerly referred to as SOBRA) was defined to be women having one of a defined set of Medicaid eligibility categories (MM P, MM T, MU, MRMP, and MRMT). All women not meeting the criteria for the No Cash Assistance group were categorized as Cash Assistance. Each maternal and infant health status indicator is expressed as a rate (either percent or average as appropriate) for statewide total and each subgroup (Medicaid vs. Non-Medicaid and Cash Assistance vs. No Cash Assistance).

Interpregnancy Interval

The interpregnancy interval was calculated by subtracting the date of the most recent termination of pregnancy or the date of the last birth from the last menstrual date as reported on the birth certificate. Deliveries with a calculated negative value for the interpregnancy interval were omitted from the Average Interpregnancy Interval calculation. An analysis of the data showed that the majority of the cases with interpregnancy intervals between zero and four months were to women who had a premature termination of the previous pregnancy.

Defining the pregnancy window

Conception date was set equal to the first non-missing value among

(1) date of last menses as given on the birth certificate
(2) date of last menses as given on the Healthy Start Prenatal Screen
(3) date of birth minus 7*estimated gestational age (in weeks) as given on the birth certificate
If none of the above quantities were non-missing, then the following value was used for conception date:

(4) date of birth minus 280 days

Table 1 summarizes the four different methods for determining conception date.

<table>
<thead>
<tr>
<th>Year</th>
<th>LMP as on Birth Certificate</th>
<th>LMP as on Healthy Start Screen</th>
<th>Clinical Estimate as on Birth Certificate</th>
<th>270 Days Prior to Delivery</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>170,719 (95.52%)</td>
<td>2,327 (1.30%)</td>
<td>3,922 (2.19%)</td>
<td>1,750 (0.98%)</td>
</tr>
<tr>
<td>1997</td>
<td>171,202 (94.53%)</td>
<td>2,979 (1.64%)</td>
<td>5,161 (2.85%)</td>
<td>1,759 (0.97%)</td>
</tr>
<tr>
<td>1998</td>
<td>171,311 (93.48%)</td>
<td>3,737 (2.04%)</td>
<td>6,443 (3.52%)</td>
<td>1,776 (0.97%)</td>
</tr>
<tr>
<td>1999</td>
<td>170,968 (93.51%)</td>
<td>3,632 (1.99%)</td>
<td>6,467 (3.54%)</td>
<td>1,760 (0.96%)</td>
</tr>
<tr>
<td>2000</td>
<td>175,725 (93.86%)</td>
<td>3,775 (2.02%)</td>
<td>5,938 (3.17%)</td>
<td>1,784 (0.95%)</td>
</tr>
</tbody>
</table>
Quality Assurances Issues

*Potential Discrepancies with Other Data Sets 1: Birth Vital Statistics*

The statistics within this report represent deliveries to women identified by Medicaid status. Data on women with a missing identifier (social security number), as well as second and subsequent births of a multiple birth, are omitted. Table 2 illustrates the percentage of records omitted from the report by reason of omission. This table does not reflect non-Florida resident data omissions.

Table 2

**SUMMARY STATISTICS FOR FLORIDA BIRTHS**

**DATA OMISSIONS: 1996 – 2000**

<table>
<thead>
<tr>
<th>YEAR</th>
<th>TOTAL FLORIDA BIRTHS</th>
<th>OMITTED: MISSING SSN</th>
<th>OMITTED: MULTIPLE BIRTHS</th>
<th>OMITTED: DUPLICATE BIRTH CERTIFICATE</th>
<th>TOTAL FLORIDA DELIVERIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>189,055</td>
<td>7,860 (4.2%)</td>
<td>2,477 (1.3%)</td>
<td>33 (0.02%)</td>
<td>178,430 (94.4%)</td>
</tr>
<tr>
<td>1997</td>
<td>191,994</td>
<td>8,235 (4.3%)</td>
<td>2,658 (1.4%)</td>
<td>14 (0.01%)</td>
<td>181,087 (94.3%)</td>
</tr>
<tr>
<td>1998</td>
<td>195,344</td>
<td>9,197 (4.7%)</td>
<td>2,880 (1.5%)</td>
<td>5 (0.00%)</td>
<td>183,262 (93.8%)</td>
</tr>
<tr>
<td>1999</td>
<td>196,669</td>
<td>10,950 (5.6%)</td>
<td>2,992 (1.5%)</td>
<td>27 (0.00%)</td>
<td>182,800 (92.9%)</td>
</tr>
<tr>
<td>2000</td>
<td>203,372</td>
<td>13,555 (6.7%)</td>
<td>2,955 (1.5%)</td>
<td>10 (0.00%)</td>
<td>187,212 (92.1%)</td>
</tr>
</tbody>
</table>

The omission of women with missing social security numbers from our tabulated rates presents a potential selection bias in favor of more optimal outcomes. This excluded group has poorer birth outcomes. It may contain recent and illegal immigrants who fall outside the health care system. In subsequent editions of this report, we hope to
distinguish this group of high-risk pregnant women and analyze their birth outcomes separately. Future tables will show birth outcomes by Medicaid status for each of the last five years and for the following groups:

1. Total number of pregnancies in Florida that resulted in a liveborn delivery
2. Singleton deliveries to mothers with social security numbers
3. Singleton deliveries to mothers without social security numbers
4. Multiple birth deliveries to mothers with social security numbers
5. Multiple birth deliveries to mothers without social security numbers

_Potential Discrepancies with Other Data Sets 2: Healthy Start_  

The three tables devoted to Healthy Start Prenatal Risks Screens are based on live births only and do not include terminations (fetal deaths, stillborns, spontaneous or elective abortion). Therefore, women, for example, who underwent a late term abortion, could have received Medicaid and Healthy Start Prenatal Risk Screening but such women would not be represented in this report’s tables.

The inherited program code previously used to merge Healthy Start Prenatal Screens to Vital Statistics produced a significant number of bad matches. This situation occurred because the methods used to match the two datasets lacked sufficient discriminating capacity. For the 2002 edition of the annual report, the merging programs were rewritten and improved through the use of more matching variables. Additionally, thousands of weakly matched records were checked by hand for accuracy. The result is that the 2002 Maternal Infant Health Status Indicators Reports contains roughly two thousand additional Healthy Start records compared to previous editions of this annual report.

As a result of reviewing previous editions of the Health Status Indicators Report with the Medicaid Program Office, it was decided to exclude two Healthy Start Indicators related to consent from the 2002 edition. Information about Healthy Start Prenatal Screen: Consent of High Risk and Healthy Start Prenatal Screen: Other Based Program Consent did not add anything additional or meaningful that was not already contained in the three other tables devoted to Healthy Start prenatal screening.
**Divergence from earlier Health Status Indicator Reports**

Infant death vital records were previously merged to birth vital records using the Department of Health’s links contained within the birth vital data. Some of these matches were incorrect and some were missing. Since there are relatively few infant deaths, a handful of mismatches can cause large errors. New programs were created to merge birth and death data. Bad matches produced by our merge were removed by hand and discrepancies between DOH and our matches were checked by hand.

Problems with previous use of Medicaid data were uncovered this year. The problems fall into two categories: (1) records for a specific Medicaid ID whose coverage periods overlap (i.e. there is a period of time that is covered by more than one record), and (2) records for a person exist under more than one Medicaid ID number. Records under more than one Medicaid ID may also have the problem of overlapping coverage periods as well. Revisions to the program code were undertaken in 2001-02 to reduce overlap and duplication of records. An attempt has been made to deal with problem (1) in two ways. First, the beginning and ending coverage dates for each Medicaid record were truncated (if necessary) to the boundary of the file’s coverage period. For example, records in the 1999 file with begin dates prior to 1999, would have their begin dates set to Jan 01, 1999, and records with end dates post-1999 would be set to Dec 31, 1999. This truncation ensures that no overlap of time periods occurs between records in different files. Second, we tried to deal with overlap within files. This problem was handled completely in the Eligibility files, as the amount of overlap was minimal and the decisions as to which records to remove were straightforward. The HMO and Medipass files, on the other hand, had a large degree of overlap and it was often unclear which record(s) should be removed, and hence overlapping records within these files still exist. Further conferring with the Medicaid office will be needed to decide how best to fix this problem.

The result is that this year’s Health Status Indicators Report gives numbers for the 2000 year that were generated by a corrected program code but leaves intact numbers for 1996 to 1999 as they appeared in previous editions of the report. With regard to the Medicaid Managed Care Report for 1999-2000, numbers from the previous year’s report
were left intact for the All Florida and Medicaid Programs tables. Numbers in the third table, Individual Medicaid HMOs, for the 1999 column do not correspond to those reported in last year’s report. It was discovered in 2001 that File 28 included records that had been left Medicaid eligibility open beyond the pregnancy window. When these records were closed and overlaps eliminated (as described in the preceding paragraph), assignment to individual HMOs were significantly altered. Thus, the numbers presented in the Individual Medicaid HMOs tables of earlier Medicaid Managed Care Reports should not be considered accurate.

**Index of Inadequate Prenatal Care**

The author of the Adequacy of Prenatal Care Utilization (APNCU) Index: (Kotelchuck, M. 1994) acknowledged three limitations of this measure:

- The APNCU Index does not measure the adequacy of the content of prenatal care; rather it measures the utilization of prenatal care.
- The APNCU Index is only as accurate as the data (the birth certificate) used to calculate it; inaccuracies in birth certificate data, particularly for prenatal care and gestational age, have been well documented.
- The APNCU Index does not adjust for the risk conditions of the mother; rather it is based on the ACOG recommendations for women with uncomplicated pregnancies. As a result, the APNCU Index produces a conservative estimate of prenatal care utilization because it underestimates the true need for prenatal care.