ATTACHMENT H.

2006 VALIDATION PROCEDURES

2006 Validation Procedures

For the 2006 validation, instead of selecting a stratified sample of cycles, systematic sampling will be used with live-birth status being the primary sorting variable for the sampling frame of treatment cycles. This will produce a sample of cycles that has the same proportion of live births and non-live births as is present in the set of all cycles for the ART program, but will not produce different sampling weights for different cycles within a program. Hence, unweighted means can be used to calculate ART program-level averages.

The following procedure will be used to select ART programs for the 2006 validation sample:

- Create an initial ART program sampling frame by including all programs that submitted ART data for 2006.
- Delete from the sampling frame those programs that were validated in 2005. Let N denote the number of ART programs in the sampling frame after deletion.
- Sort the sampling frame by an ART program's annual number of ART cycles for all age groups in 2006.
- Assign a MOS of 1.0 to each program in the sampling frame. Consequently, each ART program in the sampling frame will have the same probability of selection.
- Then use the systematic sampling procedure to select 35 clinics. The sampling interval, denoted I, will be $I = N/35$, and the probability of an ART program being selected for validation will be equal to $1/I = 35/N$. For example, if N=400, the sampling interval will be I = 400/35 = 11.4 and the probability of selection will be 1/I = 1/11.4 = 0.088.

By sorting the sampling frame by ART program size (i.e. annual number of ART cycles) and then selecting a systematic sample of programs, it will not be possible to select a "bad" sample of all large programs or all small ART programs. Moreover, the average number of annual ART cycles per program for the selected sample will be very close to the average number of annual ART cycles per program for the entire sampling frame.

Each ART program selected for validation will be reviewed as follows:

- Validation of all embryo-banking cycles,
- Full validation of a sample of treatment cycles (i.e. non-banking cycles),
- Partial validation of all live-birth cycles that are not fully validated.

To calculate discrepancy rates, one must associate the collected validation data with the appropriate sample of cycles. There are three different samples of cycles:

- Sample of embryo banking cycles. This is a one-stage cluster sample—an ART program is the cluster, and a program's entire set of embryo banking cycles is the set of cluster elements. The data obtained from the review of embryo-banking cycles are associated with this sample.
- Sample of fully-validated treatment cycles. This is a two-stage sample—an ART clinic is the first-stage sample unit, and a treatment cycle is the second-stage sample unit. Data obtained from fully-validated cycles are associated with this sample.
- Sample of live-births. This is a one-stage cluster sample—an ART program is the cluster, and a program's entire set of live-birth cycles is the set of cluster elements. Data obtained from partially-validated cycles and corresponding data obtained from fully-validated live-birth cycles are associated with this sample.

SELECTION OF CLINICS FOR VALIDATION OF 2006 DATA

Specifics of Selection Process

Of the 426 clinics submitting 2006 data, 30 clinics had been selected for 2005 data validation. These clinics were removed from the list of clinics subjected to sampling for validation of their 2006 data. The remaining 396 clinics were eligible for 2006 validation sampling, from which 35 clinics were sampled after being sorted by the total number of ART cycles within each clinic.

Clinics Selected for Validation of 2006 Data

(The names of the clinics have been concealed to protect their identity).

Clinic identifier	Number of cycles	Number of births
1	More than 500 cycles	Over 200 births
2	More than 500 cycles	Over 200 births
3	More than 500 cycles	Over 200 births
4	More than 500 cycles	Over 200 births
5	More than 500 cycles	51-100 births
6	More than 500 cycles	101-200 births
7	251-500 cycles	101-200 births
8	251-500 cycles	101-200 births
9	251-500 cycles	51-100 births
10	251-500 cycles	101-200 births
11	251-500 cycles	101-200 births
12	251-500 cycles	101-200 births
13	251-500 cycles	1-50 births
14	101-250 cycles	51-100 births
15	101-250 cycles	51-100 births
16	101-250 cycles	1-50 births
17	101-250 cycles	51-100 births
18	101-250 cycles	1-50 births
19	101-250 cycles	1-50 births
20	101-250 cycles	1-50 births
21	101-250 cycles	1-50 births
22	101-250 cycles	1-50 births
23	101-250 cycles	1-50 births
24	101-250 cycles	1-50 births
25	101-250 cycles	1-50 births
26	1-100 cycles	1-50 births
27	1-100 cycles	1-50 births
28	1-100 cycles	1-50 births
29	1-100 cycles	1-50 births
30	1-100 cycles	1-50 births
31	1-100 cycles	1-50 births
32	1-100 cycles	1-50 births
33	1-100 cycles	1-50 births
34	1-100 cycles	1-50 births
35	1-100 cycles	1-50 births

List of Variables for Full and Partial Validation

Full Validation Variables:

- Patient date of birth
- Cycle start date
- Any additional ART cycles for this patient started in 2006
- Patient diagnosis (i.e., reasons for ART)
- Oocyte or embryo source (e.g., patient, donor)
- Oocyte or embryo state (e.g., fresh, frozen)
- Transfer type (e.g., IVF, GIFT)
- Cancelled cycle notations if applicable
- Whether or not ICSI was performed during this cycle
- Transfer date
- Total number of embryos or oocytes transferred
- Outcome of treatment (e.g., biochemical only, clinical uterine gestation, ectopic)
- Ultrasound with maximum number of fetal hearts detected
- Outcome of pregnancy (e.g., live birth, spontaneous abortion)
- Information source for pregnancy outcome (e.g., patient, hospital)
- Number of infants born live
- Number of infants stillborn

Partial Validation Variables:

- Patient date of birth
- Cycle start date
- Date of infant birth
- Number of infants born live
- Number of infants stillborn
- Information source for pregnancy outcome (e.g. patient, hospital)