Appendixes

Study Procedures

QUESTIONNAIRE DEVELOPMENT

A survey questionnaire was developed by the contractor- in-concert with OTA according to the detailed research objectives set forth by OTA. OTA staff, along with the OTA advisory panel and specially selected outside experts, reviewed the draft questionnaire. A final pretest version of the questionnaire incorporated the suggestions and criticisms of the advisory panel and outside experts.

The survey instrument was pretested among 30 physicians in April 1987. The findings of the pretest concerning areas of difficulty or confusion for the respondent were used to revise the questionnaire. This revised questionnaire was approved by OTA on June 5, 1987, as the final version of the instrument. An amended version of the survey instrument was used for the survey of sperm bank practice.

SAMPLED POPULATIONS

Cross-Sectional Sample

In order to generate sample estimates that could be projected to the total population of U.S. physicians who conduct artificial insemination, a national cross-sectional sample was drawn from the universe of currently practicing physicians whose primary specialty was likely to include some infertility therapy, based on the American Medical Association (AMA) physician listings. The target population was primary care physicians likely to become involved in infertility therapy – those in general practice and family practice or in reproductive care specialties (gynecology, obstetrics/gynecology, and urology). These specialties could be sampled from the AMA sampling frame, which allows classification according to both primary and secondary specialties. The sampling frame included both office-based and hospital-based physicians; however, interns and residents were excluded.

A proportionate sampling of the population (see table A-1) would have led to relatively small sample sizes for some specialties most likely to treat fertility problems. Hence, it was decided to

sample the four specialties disproportionately, to yield 1,600 cases for the cross-sectional sample of physicians, of whom 1,575 were actually sampled.

Table A-1. - Distribution of Physicians by Specialty

Specialty	Population	Proportion	Sample size
General practice	25,807	242	388
Family practice .,	43,221	40.6	650
Gynecology/obstetrics	28,511	268	428
Urology	8,944	84	134
Total ...	106,483		1,600

SOURCE American Medical Association Physician Listings, 1987

Fertility Society Sample

Given the anticipated low physician involvement in artificial insemination and fertility treatment, this cross-sectional approach was not expected to yield a large enough sample of practitioners to permit detailed analysis. A second sampling frame was therefore constructed from the membership lists of two national professional societies, the American Fertility Society (AFS) and the American Society of Andrology. The memberships of the two organizations

are currently estimated at 11,000 and 1,000, respectively.

A prior screening of the AFS membership for fertility practices had been conducted by the association between October 1984 and March 1985. A total of 8,500 survey forms had been mailed to members and 3,200 had been returned. Among those returned, 2,736 physicians reported some practice of artificial insemination – either by donor (AID) or by husband (AIH) or both. This prescreened sample of AFS members who had identified themselves as providing artificial insemination services of any kind was used as part of the sampling frame of fertility specialists. Since no such prescreening information was available for the Andrology Society membership, the full membership list was used as the basic sampling frame.

The total size of the sample drawn of fertility specialists was 1,213. This included 1,000 from the AFS prescreened sample and 213 from the Andrology Society lists.

Sperm Bank Sample

A list of 30 separately owned or operated commercial sperm banks in the United States was developed by OTA on the basis of most current lists from the American Fertility Society and the American Association of Tissue Banks. All sperm banks were contacted with the same double mailing approach used for the physician survey. Fifteen sperm banks responded, a response rate of 50 percent. Some responses were somewhat incomplete.

SAMPLING METHOD

For all three physician samples, selection of sample within stratum was by simple random sampling, a method of selecting n units out of the N such that everyone of the distinct elements has an equal chance of being drawn. A simple random sample is drawn sequentially in practice. At any point in the draw there must be an equal chance of selection for any element in the population not already drawn.

Simple random sampling has the distinct advantage of reducing the variance of sample estimates, under most circumstances, compared with the alternative of stratified cluster sampling. Statistical formulas for specifying the sampling precision associated with particular sample sizes are based upon the assumption of simple random sampling. Cluster samples introduce a design effect into these calculations, which normally increases the expected sampling error relative to that which would have been obtained.

Simple random sampling is done by systematically selecting every "ith" person in the sampling universe. In this case the "ith" refers to a constant interval, which is determined by the formula: $i = N/n$, where N is the number of elements in the population and n is the desired number of elements in the sample. The elements in the sampling universe are listed in a random order. A computer-generated random number is used to select an initial number between 1 and i to establish a random start. The constant interval "i" is sequentially accumulated until all sampled elements have been designated. This procedure can be demonstrated to be statistically identical to the method whereby individual elements are selected at random without replacement from the population.

FIELD PROCEDURES

The field procedures used in this study were designed to produce an unbiased sample of physicians from the two sampling frames. These procedures included:

- an advance letter sent to all sample respondents indicating that the questionnaire would follow,
- a first mailing of the questionnaire with cover letter,
- a followup letter to individuals whose replies were not received within 4 weeks of the first mailing,
- a second questionnaire mailing approximately 1 week after the followup letter, and
- a telephone followup of nonrespondents among a predesignated 20 percent subset of

the sample to find out why the person had not responded.

The survey was conducted between June and August 1987. The schedule for the study did not permit a third mailing, which would have required another month for data collection. The resources available to the study did not permit the use of incentives to increase response rate. Similarly, a telephone followup with all nonrespondents was not possible within the resource limitations of the study.

PARTICIPATION RATES

These procedures yielded a total of 1,558 questionnaires completed and returned by an adjusted sample of 2,569 physicians. The overall response rate was 61 percent: 692 surveys out of an adjusted sample of 1,098 (63 percent) for the members of the fertility society sample, and 866 surveys out of an adjusted sample of 1,471 (59 percent) among respondents in the cross-sectional sample. Broken down by fertility society, 596 surveys out of an adjusted sample of 960 were completed among respondents in the AFS sample (62 percent), and 96 surveys out of an adjusted sample of 138 (70 percent) were completed among respondents in the American Society of Andrology sample (table A-2).

The survey was designed to permit a systematic effort to define the sources of nonresponse on a limited basis. A random sample of one-fifth of the total was predesignated for followup before the field work began. As the field period ended, all outstanding cases from this sample were contacted in an attempt to learn why they were not responding. No attempt was made to administer the survey over the telephone or to collect demographics; rather, the telephone followup aimed to identify any source of systematic bias in the achieved sample.

Roughly 35 percent of the contacted nonrespondents reported that they had already completed the survey and just recently returned it by mail, that they intended to reply, or that they

were in the process of replying as of the date of the telephone followup call. It seems that additional prodding, in the form of the followup call, actually improves overall response rates, as indicated by this high "will reply" response (table A-3). This also suggests that a third-wave mailing might have produced a higher response rate, possibly 70 percent.

A large percentage of the nonresponses involved those who were unreachable at the time of the followup call. Almost 25 percent of the nonresponse sample were on vacation, not at home, or otherwise unreachable at the time.

Another important cause for nonresponse is straightforward refusal to participate in the survey. Roughly 15 percent of the non response sample refused to participate for a variety of reasons, **most** of which involved the length and/or complexity of the survey instrument or the respondents' actual time available to complete the survey. The construction of the survey did not seem to be a factor in nonresponse, with only a few respondents citing this as a reason for refusal. A few physicians cited issues of privacy, lack of incentive or benefit in completing the survey, or a policy of refusal to participate in surveys.

Other sources for nonresponse include those "ineligible" for a variety of reasons, such as those no longer practicing, those no longer at the prestated location who left no forwarding address, and those deceased. Ineligibility accounted for

Table A-2.-Sample Disposition

	Cross-section	AFS	Andrologists
Total sampled	1,575	1,000	213
Ineligible	76	15	47
No infertility work	26	5	2
Not a practicing MD	9	1	43
Moved out of U.S.	0	1	0
Deceased	2	1	1
Retired	37	5	0
Other misc.	2	2	1
Bad address/no forwarding	28	25	28
Adjusted sample	1,471	960	138
No return	584	335	38
Remails	3	62	0
No reply	578	232	38
On vacation	3	0	0
New address from AFS	0	32	0
Not listed by AFS	0	9	0
Refused	21	29	4
Completes received	866	596	96
>4 inseminations	37	361	24
0-3 inseminations	819	231	71
Unspecified	10	4	1
Completes used in analysis	827	569	77
>4 inseminations	36	346	21
Others	791	223	56
Completes received too late	39	27	19

SOURCE: Office of Technology Assessment 1988

30 percent of the nonresponse for the preselected group derived from the Andrology Society sample, but accounted for no more than 8 Percent of those preselected respondents from the AFS and the cross-sectional samples. Overall, the followup contact did not reveal any underlying problem of sample bias among nonrespondents.

DIFFERENTIAL RESPONSE RATES AND SAMPLE WEIGHTING

Differential response rates by physician speciality and professional society membership can produce some sample distortion from true population distribution. Further, for the cross-sectional sample, the sample drawn was not representative of true population distribution for the four sampled fields of medical practice. To correct for such biases, the distribution of the achieved sample was compared with the actual distribution by specialty and society for these groups, and sample weights were applied to correct for differences.

PRECISION OF SAMPLE ESTIMATES

The objective of the sampling and field procedures is to produce an unbiased sample of the target population, one that shares the same properties and characteristics of the total population from which it is drawn, subject to a certain level of sampling error. This means that with a properly drawn sample, statements can be made about the properties and characteristics of the

Table A-3.– Phone Results of Telephone Followup Among Outstanding Cases In Predesignated Followup Sample

	Cross-section	AFS	Andrologists
Total sampled	316	201	43
Outstanding cases	144	107	14
Survey is in the mail	9	9	1
Will reply/in the process	29	24	2
Remails	31	11	2
Remail to new address..	9	5	1
Need new copy (same address)	13	1	0
Have not received (same address)	9	5	1
Ineligible	10	2	4
Screen out..	9	0	2
No longer at location	1	2	1
Deceased	0	0	1
Not reached	22	24	3
On vacation.	3	9	0
Call back for doctor	12	14	2
No answer/busy	2	1	1
Left message on answering machine	5	0	0
Have not called	30	21	0
Insufficient time since last remail of survey	0	11	0
Wrong telephone number	1	1	0
Tel. number not listed	25	9	0
Tel. number unpublished	4	0	0
Refusals			
Survey is too long/complex	2	5	0
Do not have time.....	4	5	0
Survey is poorly constructed	0	2	1
No benefit in doing survey	1	0	0
Privacy/no reason for info requested	1	0	0
Mad at government/lower taxes, etc.	2	0	0
Doctors should be paid	1	0	0
Do not do surveys	1	0	0
Other/unspecified reason	1	3	0
Request remail -too late to comply	0	0	0
Duplicate/same clinic (won't reply)	0	1	1
Completed surveys	30	22	3
Adjusted sample	104	63	9

SOURCE: Office of Technology Assessment, 1988

total population within certain specified limits of certainty and sampling variability.

The expected sampling error for sample estimates of population proportions, using simple random sampling without replacement, is calculated by the following formula:

$$\mathrm{var}(x) = (z)\left(\frac{\sqrt{p(q)}}{n-1}\right)$$

where:

var $(x) =$ the expected sampling error of the mean of some variable, expressed as a proportion;

$P =$ some proportion of the sample displaying a certain characteristic or attribute;

$q =$ $(1-p)$;

$z =$ the standardized normal variable, given a specified confidence level; and

$n =$ the size of the sample.

The maximum expected sampling error at the 95 percent confidence level (i.e., in 95 out of 100 repeated samples) for a total physician sample of 827 (the cross-section) is +/- 3.4 percentage points. It should be noted that the maximum sampling error is based upon a certain response distribution (i.e., a 50/50 split). The expected sampling error is less for other response distributions with a sample size of 827. For example, the estimate that 9.6 percent of a sample of 827 physicians have accepted patients for artificial insemination in the past year is subject to an expected sampling error of +/- 2.0 at the 95-percent confidence level. However, as sample size declines, as in the case of subsamples of the total physician sample, so too will the expected sampling precision of the estimates. Table A-4 presents the expected size of the sampling error for specified sample sizes of 1,500 and less, at different response distributions on a categorical variable. This table may be used to project the estimated precision of sampling estimates for the total sample and naturally occurring subsets of the sample, e.g., particular medical specialties, number of years in practice, and so on.

Table A4.-Expected Sampling Error (plus or minus) at the 95 Percent Confidence Level (simple random sample)

Size of sample or subsample	Percentage of the sample or subsample giving a certain response or displaying a certain characteristic for percentages near: [a]				
	10 or 90	20 or 80	30 or 70	40 or 60	50
1,500	1.5	2.0	2.3	2.5	2.5
1,300	1.6	2.2	2.5	2.7	2.7
1,200	1.7	2.3	2.6	2.8	2.8
1,100	1.8	2.4	2.7	2.9	3.0
1,000	1,9	2.5	2.8	3.0	3.1
900, ...	2.0	2.6	3.0	3.2	3.3
800	2.1	2.8	3.2	3.4	3.5
700	2.2	3.0	3.4	3.6	3.7
600	2.4	3.2	3.7	3.9	4.0
500	2.6	3.5	4.0	4.3	4.4
400	2.9	3.9	4.5	4.8	4.9
300	3.4	4.5	5.2	5.6	5.7
200	4.2	5.6	6.4	6.8	6.9
150	4.8	6.4	7.4	7.9	8.0
100	5.9	7.9	9.0	9.7	9.8
75	6.8	9.1	10.4	11,2	11.4
50	8.4	11.2	12.8	13.7	14.0

a Entries are expressed a percentage points (+ or -).

SOURCE: Office of Technology Assessment, 1988

ESTIMATING STATISTICAL SIGNIFICANCE

The estimates of sampling precision presented in the preceding section yield confidence bands around the sample estimates, within which the true population value should lie. This type of sampling estimate is appropriate when the goal of the research is to estimate a population distribution. When the goal is to compare the survey responses between two or more populations, however (i.e., to determine whether the characteristics of two populations are different), it is necessary to consider whether differences observed between the samples are statistically significant (i.e., beyond the expected limits of sampling error for both sample estimates).

To test this, a rather simple calculation can be made. Call the total sampling error (i.e., var (x) in the previous formula) of the first sample s1 and the total sampling error of the second sample s2. The sampling error of the difference between these estimates is sd, which is calculated as:

$$sd = \sqrt{s1^2 + s2^2}$$

Any difference between observed proportions that exceeds sd is a statistically significant difference at the specified confidence internal. Note that this technique is mathematically equivalent to generating standardized tests of the difference between proportions.

An illustration of the pooled sampling error between subsamples for various sizes is presented in table A-5.

Table A-5.–Pooled Sampling Error Expressed As Percentages for Given Sample Sizes
(assuming p=q)

Sample size	100	200	300	400	500	600	700	800	900	1000	2000
2,000	10.0	7.2	6.1	54	4.9	4.6	43	4.1	4.0	3.8	3.1
1,000	10.3	7.6	6.5	5.8	54	51	49	4.7	4.5	4.4	
900	10.3	7.6	6.6	5.9	5.5	5.2	50	48	4.7		
800	10.4	7.7	6.7	60	5.6	53	5.1	5.0			
700	10.5	7.8	6.8	6.1	5.8	5.4	52				
600	10.6	8.0	7.0	6.3	59	5.7					
500	10.7	82	72	6.6	6.2						
400	11.0	85	75	69							
300	11.3	9.0	8.1								
200	12.0	9.8									
100	13.9										
Sample size	100	200	300	400	500	600	700	800	900	1000	2000