American Journal of Epidemiology

Vol. 161, No. 10
Printed in U.S.A.
DOI: 10.1093/aje/kwi120

Validity of Self-reported Causes of Subfertility

Evelien J. de Boer[1,2], Isolde den Tonkelaar[2], Curt W. Burger[3], and Flora E. van Leeuwen[1] for the OMEGA Project Group

[1] Department of Epidemiology, Netherlands Cancer Institute, Amsterdam, the Netherlands.
[2] International Health Foundation, Utrecht, the Netherlands.
[3] Department of Gynecology and Obstetrics, Erasmus Medical Center, Rotterdam, the Netherlands.

Received for publication January 8, 2004; accepted for publication January 5, 2005.

The authors assessed the accuracy of cause(s) of subfertility as reported by women in a self-administered questionnaire in comparison with medical record information, in a nationwide cohort study of women receiving in vitro fertilization treatment in the Netherlands ($n = 9{,}164$) between 1983 and 1995. Validity was expressed as sensitivity and specificity, and reliability was expressed by the kappa statistic and overall agreement between self-reports and medical records for various subfertility categories. The sensitivity for subfertility attributed to tubal, male, hormonal, cervical, uterine, and idiopathic factors and for endometriosis was 84%, 78%, 65%, 40%, 46%, 59%, and 83%, respectively. The corresponding kappas were 0.79, 0.71, 0.38, 0.34, 0.13, 0.50, and 0.52, respectively. For 54% of all women who reported two or more causes of subfertility, the medical record revealed only one major factor. Conversely, for 43% of all women whose subfertility was attributed to two or more major factors in the record, only one factor was reported by the women. Older age at the time of filling out the questionnaire, low educational level, long duration of subfertility, and pre–in vitro fertilization treatment were associated with less accurate reporting. The results indicate that the validity of self-reports for tubal and male subfertility is satisfactory. For unexplained subfertility, the validity is moderate; for other causes of subfertility and when two causes of subfertility play a role, the validity is low.

cohort studies; fertility; fertilization in vitro; medical records; questionnaires; reproductive techniques, assisted; sensitivity and specificity; women

Abbreviation: OMEGA, Ovariële hyperstiMulatie En Gynaecologische Aandoeningen.

Information regarding the cause of subfertility is often needed in epidemiologic studies assessing risk factors for hormone-related cancers and gynecologic disorders (1–5). Although the medical record is the most accurate source for information regarding cause(s) of subfertility, the huge investment needed to gain access to and abstract the medical record often prevents researchers from embarking on such an enterprise. When information from the medical records is not easily available, information on the cause of subfertility is obtained from the women themselves by mailed questionnaires or personal interviews. Although some studies examining the association between the cause of subfertility and breast or ovarian cancer risk are based upon medical record information (2, 6), many studies are based upon self-reported data only (1, 3, 7–10).

To our knowledge, no reports are available on the validity and reliability of self-reported data on the cause of subfertility. Some studies evaluated the recall of menstrual and reproductive factors and reported a satisfactory-to-good recall for age at menarche and age at menopause, pregnancy-related events, and birth characteristics (11–15). Accuracy of recall was less for menstrual cycle characteristics (11, 14, 16).

In a large Dutch cohort of women who underwent in vitro fertilization, we assessed the validity of self-reported causes of subfertility in comparison with medical record

Correspondence to Flora E. van Leeuwen, Department of Epidemiology, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands (e-mail: f.v.leeuwen@nki.nl).

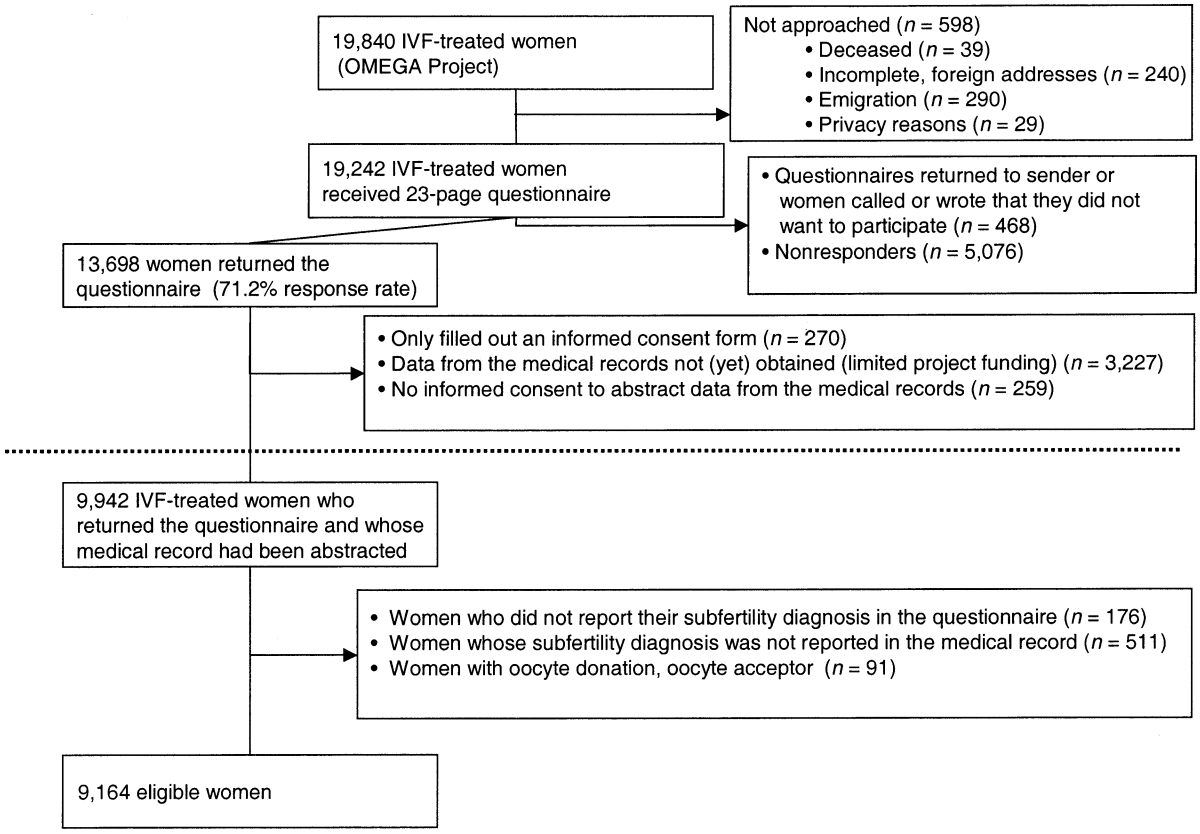

FIGURE 1. Description of the recruitment of eligible women from the OMEGA study population, Amsterdam, the Netherlands, 1983–1995. OMEGA, Ovariële hyperstiMulatie En Gynaecologische Aandoeningen; IVF, in vitro fertilization.

information. Detailed information on the causes of subfertility was available from two sources, that is, mailed questionnaires from the study participants and medical records that were abstracted by trained research assistants. With the advent of assisted reproductive technology, subfertile couples undergo a variety of diagnostic procedures prior to the start of subfertility treatment. The results of these tests may point to a single cause of subfertility but may also indicate two or even more contributing factors. Both the women and the trained research assistants in our study could report more than one cause of subfertility. Consequently, we could also investigate differences in validity of recall between women with only one and those with two causes of subfertility. Furthermore, we were able to examine the role of several covariates on the accuracy of reporting.

MATERIALS AND METHODS

Study population and study procedures

The Ovariële hyperstiMulatie En Gynaecologische Aandoeningen (OMEGA) study population, procedures, and data collection methods have been described previously in detail (5, 17, 18). Briefly, subjects are participants in a nationwide cohort study of 19,242 women who received treatment by in vitro fertilization in the Netherlands between 1983 and 1995. The main purpose of this study is to examine the risk of hormone-related cancers in women receiving in vitro fertilization treatment. All institutional ethics committees of the participating clinics approved the study procedures.

Between 1997 and 2000, all women received a mailed questionnaire on reproductive history, history of subfertility (treatment), the use of exogenous hormones (fertility drugs, oral contraceptives, and hormone replacement therapy), and various lifestyle factors; women also received an informed consent form for data abstraction from the medical records. A total of 13,698 women returned the questionnaire (response rate: 71.2 percent). The upper part of figure 1 displays a graphical presentation of the OMEGA study population.

Women were eligible for the present study if they gave permission for data abstraction from the medical files and their medical file had been abstracted ($n = 9,942$). For 3,227 women who were treated by in vitro fertilization and who returned the questionnaire, data from the medical files could not yet be obtained. Since this was because of limited project funding resulting in a random sample of records not being completed, it is unlikely that validity of recall among

these women would have been different. After exclusion of women with an unknown subfertility diagnosis ($n = 687$) and women who underwent oocyte donation ($n = 91$), 9,164 women remained in the analyses (see lower part of figure 1).

Data collection and methods of analyses

In the questionnaire, women were asked to report their cause(s) of subfertility according to 12 categories, that is, tubal obstruction, factors related to the male partner, endometriosis, ovulation disorders, other hormonal disorders, premature ovarian failure, polycystic ovary syndrome, uterine abnormalities, cervical disorders, factors related to vaginal disorders, unexplained subfertility (despite full subfertility assessment, no particular cause of subfertility found), and other factors. Women could report more than one cause of subfertility (table 1). For the present study, the self-reported causes of subfertility were classified into eight categories: tubal, male, ovarian (including ovulation disorders, polycystic ovary syndrome, premature ovarian failure, and hormonal disorders related to ovarian function), cervical disorders, uterine abnormalities, endometriosis, unexplained, and other (including vaginal disorders).

Trained research assistants abstracted the medical files using a standardized questionnaire, to obtain information on the gynecologic history, subfertility diagnosis, and subsequent in vitro fertilization treatment cycle. The subfertility diagnosis was classified as due to tubal factors, male factors, ovarian factors (including ovulation disorders, polycystic ovary syndrome, premature ovarian failure, and hormonal disorders related to ovarian function), cervical factors (including antisperm antibodies), uterine abnormalities, endometriosis, unexplained subfertility, and other. The research assistants classified each potential cause of subfertility according to the extent to which it contributed to subfertility (strongly, probably, little, or not at all). As with the questionnaire, the medical records could have listed more than one type of subfertility, with different classifications. In the present analysis, we combined the classifications ''strongly'' or ''probably'' into one classification, which was considered a major contribution to subfertility.

We assessed the agreement between self-reports and medical records, adjusting for chance agreement by use of the Cohen kappa statistic, which was estimated for various subfertility categories. Values of kappa greater than 0.75 represent excellent agreement, values of 0.40–0.75 were considered to indicate ''moderate'' agreement, and a value of less than 0.40 represents poor agreement (19).

Furthermore, we calculated the overall proportion of observed agreement (positive and negative) between self-reports and medical records. In addition, we analyzed the overall kappa value of all types of subfertility within a subgroup of women whose subfertility was (''strongly'') attributed to only one factor, as reported in both the questionnaire and medical record.

Using the medical record information as our ''gold standard,'' we determined the validity measures sensitivity and specificity and positive and negative predictive values. Sensitivity is defined as the proportion of women with a specific cause of subfertility according to medical record who correctly classified themselves with this particular cause of subfertility in the questionnaire. Specificity is the proportion of women who truly do not have a specific cause of subfertility (according to the medical record) and who correctly classified themselves as such in the questionnaire.

Logistic regression analysis was used to determine which variables were independently associated with overall agreement regarding the subfertility diagnosis reported in the questionnaire and the medical records. The variables of interest were age at completion of the questionnaire ($\leq$34, 35–37, 38–40, and $>$40 years), duration of subfertility before in vitro fertilization ($\leq$1, 2–3, 4–5, and $\geq$6 years), pre–in vitro fertilization treatment (none, fertility drug use only, and intrauterine insemination with or without fertility drug use), years since first in vitro fertilization treatment as a ''proxy'' for the period of time that had elapsed since the woman received information about the cause(s) of subfertility ($\leq$2, 3–5, 6–7, and $\geq$8 years), number of in vitro fertilization treatment cycles (one, two, three, and greater than three), educational level (low (primary school), middle (secondary school), high (college or university), and unknown), and whether or not women had at least one livebirth after in vitro fertilization treatment. All analyses were processed with SPSS, version 11.0, software (SPSS, Inc., Chicago, Illinois).

RESULTS

Table 1 presents the general characteristics of the study population. The median age at the completion of the questionnaire was 38.7 years. The median time interval between the first in vitro fertilization treatment and the completion of the questionnaire was 5.5 years. The first in vitro fertilization treatment occurred less than 2 years before in 12 percent of the women and 8 or more years before in 18 percent of the women. The median number of in vitro fertilization cycles was three, and 64 percent of the women never had a livebirth after treatment with in vitro fertilization. A total of 5,874 (64 percent) women reported that only one factor contributed to their subfertility. According to the medical record, 83 percent of all women had only one factor that ''strongly'' contributed to their overall subfertility. In 54 percent of all women who reported two or more causes of subfertility, the medical record revealed that subfertility was attributed to only one major factor. Conversely, for 43 percent of all women whose subfertility was attributed to two of more major factors according to the medical record, the women themselves reported only one cause of subfertility.

Table 2 presents the comparison between self-reports and medical records for various subfertility diagnoses. The most important causes of subfertility as reported by the women were tubal (38 percent), male (38 percent), unexplained (23 percent), endometriosis (13 percent), and hormonal (17 percent). These percentages add up to more than 100 percent because more than one cause of subfertility could be reported by the women (and abstracted from the medical records). Of all the women who attributed their subfertility to tubal factors, this cause was confirmed in the medical

TABLE 1. Characteristics of the study population, including the number of subfertility diagnoses according to self-reports in questionnaires and medical records ($n = 9,164$), Amsterdam, the Netherlands, 1983–1995

Characteristic	No. of participants	% of total
Age (years) at questionnaire		
≤34	1,939	21.2
35–37	2,112	23.0
38–40	2,250	24.6
>40	2,863	31.2
Years since first in vitro fertilization treatment*		
≤2	1,090	11.9
3–5	4,201	45.8
6–7	2,244	24.5
≥8	1,608	17.5
Duration of subfertility before in vitro fertilization (years)		
≤1	2,011	21.9
2–3	2,838	31.0
4–5	1,014	11.1
≥6	2,084	22.7
Unknown	1,217	13.3
Pre–in vitro fertilization fertility treatment		
Use of fertility drugs only	1,622	17.7
Intrauterine inseminations	2,169	23.7
None	5,373	58.6
No. of in vitro fertilization treatment cycles		
1	2,001	21.8
2	2,053	22.4
3	2,408	26.3
>3	2,668	29.1
Educational level†		
Low	2,511	27.4
Middle	4,468	48.8
High	2,002	21.8
Unknown	183	2.0
Livebirth after in vitro fertilization treatment with at least one baby born alive		
No livebirth	5,840	63.7
≥1 livebirth	3,324	36.3
No. of self-reported subfertility diagnoses		
1	5,874	64.1
2	2,393	26.1
≥3	897	9.8
No. of diagnoses "strongly" contributing to subfertility, as reported in the medical records		
0	304	3.3
1	7,612	83.1
2	1,197	13.1
≥3	51	0.6
No. of diagnoses with a major contribution to subfertility, as reported in the medical records‡		
1	6,495	70.9
2	2,380	26.0
≥3	289	3.2

* Numbers do not add up to 100% because of missing data.

† Educational level: low (primary school), middle (secondary school), high (college or university), and unknown.

‡ The classifications "strongly" or "probably" contributing to subfertility were combined into one classification, which was considered to make a major contribution to the subfertility.

record in 91 percent of all cases (positive predictive value). Conversely, of all the women whose subfertility was caused by tubal factors according to the medical record, 84 percent reported tubal subfertility in the mailed questionnaire (sensitivity). The overall agreement (positive and negative) for tubal subfertility was 90 percent, and the kappa value was 0.79, indicating excellent agreement. For male subfertility, the validity measures and agreement were just a bit lower and also indicative of accurate reporting. For both unexplained subfertility and endometriosis, the kappa values indicated "moderate" agreement, with rather low positive predictive values (65 percent and 43 percent, respectively). Despite the low kappa value and sensitivity for women whose subfertility was attributed to hormonal, cervical, or uterine factors, the observed specificity was high.

Table 3 presents the comparison between self-reports and medical records for the overall type of subfertility among the 5,874 women who reported only one cause of subfertility. Among women whose medical records revealed that only one "strong" factor contributed to their overall subfertility ($n = 5,269$, 90 percent), the kappa value for all causes of subfertility was 0.78, indicating excellent agreement. The positive predictive values for subfertility attributed to tubal, male, unexplained, endometriosis, hormonal, and cervical subfertility were 87 percent, 80 percent, 70 percent, 33 percent, 33 percent, and 36 percent, respectively. For 441 of the 1,467 women (30 percent) who reported that their subfertility was attributed to idiopathic factors only, the medical record revealed one or more specific causes to which subfertility was attributed; in 171 of these women (12 percent), subfertility was due to a male factor. Conversely, of the 1,396 women whose cause of overall subfertility was unexplained according to the record, 329 women (24 percent) themselves reported that their overall subfertility was due to male factors, tubal factors, endometriosis, hormonal factors, or cervical factors.

Since accurate recall of more than one cause of subfertility may be more difficult than recall of only one factor, we analyzed the agreement between self-reported causes of subfertility and medical record data separately for women who reported two causes of subfertility (data not shown). Of all the women who reported both tubal and male subfertility (which was the most frequent combination), these causes were confirmed in the medical record for 53 percent of the women. Sensitivity for the combination of tubal and male subfertility was only 32 percent, while sensitivity was 70 percent and 57 percent, respectively, for subfertility attributed only to tubal or only to male factors (data not shown). A quarter of all women whose subfertility was attributed to tubal and male factors in the record reported only tubal factors in the questionnaire (data not shown).

Table 4 presents the accuracy of self-reported cause of subfertility in comparison with the medical records according to several characteristics of the study population. For women aged 38 years or more at the time of filling out the questionnaire, the adjusted odds ratio for accurate recall of type of subfertility was 0.62 (95 percent confidence interval: 0.49, 0.80) compared with that for women aged less than or equal to 34 years. Furthermore, women with a high educational level were more likely to recall their type of subfertility

TABLE 2. Comparison between self-reports from mailed questionnaires and medical records for various subfertility diagnoses (n = 9,164), Amsterdam, the Netherlands, 1983–1995*

Cause of subfertility†	Self-report (+)		Self-report (−)		Validity			Reliability	
	Medical record (+)‡ (no.)	Medical record (−)‡ (no.)	Medical record (+)‡ (no.)	Medical record (−)‡ (no.)	Positive predictive value (%)	Sensitivity (%)	Specificity (%)	Agreement (%)	Kappa statistic
Tubal	3,174	324	607	5,059	91	84	94	90	0.79
Male	3,078	446	862	4,778	87	78	91	86	0.71
Unexplained§	1,352	732	929	6,151	65	59	89	82	0.50
Endometriosis	518	689	103	7,854	43	83	92	91	0.52
Hormonal	548	1,020	309	7,287	35	65	88	86	0.38
Cervical	192	352	292	8,328	35	40	96	93	0.34
Uterine	32	361	37	8,734	8	46	96	96	0.13

* Because of the absence of a uniform definition for women categorized into the so-called other cause of subfertility subgroup in both questionnaire and medical record, these women were excluded from the analysis calculating the kappa statistic.

† More than one cause of subfertility could be reported by the women and abstracted from the medical records.

‡ The classifications "strongly" or "probably" contributing to subfertility were combined into one classification, which was considered to make a major contribution to the subfertility.

§ Despite full subfertility assessment, no particular cause of subfertility was found.

accurately than were women with a low educational level. A long period of subfertility and pre–in vitro fertilization fertility treatment were associated with less accurate reporting of the cause of subfertility.

DISCUSSION

We observed that the sensitivity, positive predictive value, and agreement as expressed by kappa were excellent for tubal and male subfertility and substantially lower for other causes. For unexplained subfertility, the positive predictive value and sensitivity were moderate. Sensitivity was high for subfertility attributed to endometriosis and moderate for hormonal factors and uterine abnormalities, while the positive predictive values were low for these diagnoses. Specificity was remarkably high for all causes of subfertility.

The highly accurate self-report for tubal and male subfertility may be explained by the fact that these are widely known causes of subfertility. In the first years after the introduction of in vitro fertilization treatment in the Netherlands, the major indication for treatment with in vitro fertilization was tubal obstruction, while male subfertility became a major indication for in vitro fertilization in the mid-1980s. Tubal and male subfertility have remained the most important indications for in vitro fertilization treatment. Women might remember these causes more accurately as compared with, for example, hormonal subfertility, which is less prevalent and more difficult to understand.

For endometriosis and hormonal factors, sensitivity was high to moderate, but the positive predictive values were low. A possible explanation is that, in the era of assisted reproductive technology, couples undergo extensive fertility assessment; in the course of diagnostic work-up (which may already have started before the first in vitro fertilization treatment), it is quite possible that several subtle abnormalities are detected and communicated to the subfertile couple. Although such test results may be considered minor contributors to overall subfertility by the gynecologist, the women in our study may have remembered them and reported them in the questionnaire. This phenomenon may also explain the rather low sensitivity for "unexplained" subfertility. The distribution of subfertility causes attributed to only one factor as reported by the women in our study is more or less in line with that in the literature, which reports that 35 percent of all subfertility in couples is due to tubal factors, 35 percent to male factors, 5 percent to unusual problems, and 15 percent to ovulatory dysfunction, with 10 percent remaining unexplained (20). The percentage of unexplained subfertility in our study is higher while hormonal subfertility was less prevalent, which may be explained by the differences in source populations, the definitions used, and the differences in fertility assessment over time. In daily clinical practice, many women with hormonal subfertility are successfully treated without the necessity of in vitro fertilization treatment, explaining the lower proportion of women with hormonal subfertility in our study. Our study population was treated with in vitro fertilization between 1983 and 1994. The relatively high proportion of women diagnosed with unexplained subfertility may be due to more thorough fertility assessment in recent years.

In 54 percent of all women who reported two or more causes of subfertility, the medical record revealed that subfertility was attributed to only one major factor. Diagnostic tests in reproductive medicine may reveal subtle abnormalities and, in addition, do not always yield unambiguous results. Although the gynecologist did not consider subtle abnormalities as a major cause of subfertility, a woman might consider such abnormalities as an important cause of subfertility. Conversely, 43 percent of all the women whose subfertility was due to two or more major factors according to the medical record reported only one

TABLE 3. Comparison between self-reports and medical records for overall type of subfertility among women who reported only one cause of subfertility ($n = 5,874$), Amsterdam, the Netherlands, 1983–1995

Self-reported causes in questionnaire	No. of subfertility causes in the medical record (and classifications)										Total no.
	One cause of subfertility "strongly" contributed to the overall subfertility ($n = 5,269$)*								Two or more causes "strongly" contributed to the overall subfertility ($n = 463$)†	One or more causes "probably" contributed to the overall subfertility ($n = 142$)	
	Tubal	Male	Unexplained	Endometriosis	Hormonal	Cervical	Uterine	Other			
Tubal	1,819 (87)‡	12	64	5	3	3		3	140	35	2,084
Male	9	1,268 (80)	78		9	5		21	178	17	1,585
Unexplained§	67	171	1,026 (70)	21	25	31	1	16	60	49	1,467
Endometriosis	26	10	64	82 (33)	1			2	45	15	245
Hormonal	21	24	91	1	88 (33)	4		4	24	14	268
Cervical	3	15	32			34 (36)			7	4	95
Uterine	11	3	10			1	1 (NA¶)		2	2	30
Other	27	14	31	1	2	3		9	7	6	100

* Kappa = 0.78 (subgroup "other" not included).
† Most frequent combinations of two factors both "strongly" contributing to the overall subfertility (medical record information): (male/tubal ($n = 119$), male/unexplained ($n = 116$), tubal/unexplained ($n = 21$), and tubal/endometriosis ($n = 35$)).
‡ Numbers in parentheses, positive predictive value.
§ Despite full subfertility assessment, no particular cause of subfertility was found.
¶ NA, not applicable.

cause of subfertility themselves. In these cases, full fertility assessment revealed more than one factor classified as either "strongly" or "probably" contributing to the subfertility, but the women failed to report it in our questionnaire. The lower sensitivity for the most frequent combinations of two causes of subfertility suggests that women with two causes of subfertility are less likely to recall their subfertility accurately as compared with those who had only one cause of subfertility.

When considering the factors impacting on the accuracy of self-reports (table 4), we observed that younger age (<37 years) at questionnaire completion, a high educational level, a short period of subfertility duration (less than 2 years), and no fertility treatment prior to in vitro fertilization favorably affected the recall of cause of subfertility. Less accurate recall among women with a long duration of subfertility and pre–in vitro fertilization treatment may be explained by the higher prevalence of unexplained subfertility among these women. In our study, the time since in vitro fertilization treatment did not affect the accuracy of recall. Women who had their first in vitro fertilization treatment 8 or more years ago were able to recall their subfertility cause as accurately as did women whose treatment started 2 years ago. As pointed out above, tubal subfertility was the major indication for in vitro fertilization in the early years after its introduction. This was confirmed in our data; that is, those with a long follow-up period were more likely to have been diagnosed with tubal subfertility (54 percent). These women may have had a wish for a child for many years before the in vitro fertilization era, and they probably consulted several fertility experts while trying to achieve a pregnancy. Although our data show that the physician-patient communication must have been relatively good in the early era of assisted reproductive technology treatment, it is possible that the women treated in the first years of in vitro fertilization were more often told their subfertility diagnosis, resulting in accurate recall despite the longer time interval.

Since hardly any reports are available on the validation of the self-reported cause of subfertility, we cannot compare our results with those of the literature. A recent case-control study of ovarian cancer reported that validation of women's self-reported difficulties in conceiving (as a measure of infertility) with medical record information resulted in great difficulties. Of the 179 women for whom medical records were available, only 41 (23 percent) had their self-report of a fertility problem confirmed (21). Validation of self-reported ovulatory infertility was attempted in a randomly selected sample of 100 women participating in the Nurses' Health Study II (22). Of the 71 women who gave permission to review their medical records, only 40 medical records could be located, of which in 95 percent ovulatory infertility was confirmed by diagnostic test or specific treatment (22). Sensitivity was not assessed in this small sample of the study, however, and recall of other causes of subfertility was not examined.

When our results are interpreted, the strengths and limitations of the study design need to be considered. The strengths of our study include the large study size and the availability of detailed information on all causes of subfertility from two sources. Furthermore, the women filling

TABLE 4. Accuracy of self-reported type of subfertility according to several characteristics, among women with only one ("strong") factor contributing to the overall subfertility ($n = 5,139$), Amsterdam, the Netherlands, 1983–1995*

Characteristic	Nonagreement ($n = 821$) (no.)	Agreement ($n = 4,318$) (no.)	Univariate odds ratio	95% confidence interval	Multivariate odds ratio†	95% confidence interval
Age (years) at questionnaire						
≤34	145	966	1.0		1.0	
35–37	169	1,035	0.92	0.72, 1.17	0.91	0.71, 1.17
38–40	222	997	0.67	0.54, 0.85	0.62	0.49, 0.80
>40	285	1,320	0.70	0.56, 0.86	0.62	0.48, 0.80
Years since first in vitro fertilization treatment‡						
≤2	89	541	1.0		1.0	
3–5	385	2,000	0.86	0.67, 1.10	0.92	0.71, 1.19
6–7	215	1,007	0.77	0.59, 1.01	0.84	0.62, 1.13
>8	131	761	0.96	0.70, 1.28	0.93	0.66, 1.30
Educational level§						
Low	250	1,193	1.0		1.0	
Middle	399	2,129	1.12	0.94, 1.33	1.10	0.92, 1.32
High	151	916	1.27	1.02, 1.58	1.27	1.01, 1.60
Unknown	21	80	0.80	0.48, 1.32	0.76	0.46, 1.27
Duration of subfertility before in vitro fertilization (years)						
≤1	126	1,072	1.0		1.0	
2–3	253	1,362	0.63	0.50, 0.80	0.72	0.56, 0.91
4–5	116	440	0.45	0.34, 0.59	0.54	0.40, 0.72
≥6	202	887	0.52	0.41, 0.66	0.63	0.50, 0.84
Unknown	124	557	0.53	0.40, 0.69	0.43	0.32, 0.58
Pre–in vitro fertilization fertility treatment						
None	415	2,919	1.0		1.0	
Use of fertility drugs only	184	579	0.45	0.37, 0.54	0.60	0.49, 0.75
Intrauterine inseminations	222	820	0.53	0.44, 0.63	0.73	0.60, 0.90
No. of in vitro fertilization treatment cycles*						
1	194	1,015	1.0		1.0	
2	186	918	0.94	0.76, 1.18	0.92	0.74, 1.16
3	248	1,117	0.86	0.70, 1.06	0.85	0.68, 1.05
>3	189	1,256	1.27	1.02, 1.58	1.11	0.88, 1.40
Livebirth after in vitro fertilization treatment with at least one baby born alive						
No livebirth	538	2,637	1.0		1.0	
≥1 livebirth	283	1,681	1.21	1.04, 1.42	1.17	0.99, 1.37

* Women in the so-called other subgroup of subfertility were excluded from the analyses ($n = 30$).

† Adjusted for all the other variables; years since in vitro fertilization treatment, duration of subfertility, no. of in vitro fertilization cycles, pre–in vitro fertilization treatment, and livebirth were also adjusted for subfertility diagnosis (women were categorized into the subfertility category that was assumed to contribute most to the subfertility; four categories).

‡ $n = 5,129$.

§ Educational level: low (primary school), middle (secondary school), high (college or university), and unknown.

out the questionnaires and the research assistants abstracting the medical record were instructed that they could record more than one factor that contributed to the overall subfertility. In addition, we collected information from the medical record on the importance of all subfertility causes mentioned in the record.

A limitation of our study is, however, that in calculating the sensitivity and specificity we considered the medical record as the gold standard, thus assuming the medical records to be correct and complete. Of all the eligible women in the study population ($n = 9{,}942$), only 7 percent of the women had to be excluded because of incompleteness of the medical records for the cause of subfertility. However, some medical records that were considered by us to be complete may have been incomplete for one of the subfertility causes considered, resulting in an underestimation of overall agreement and positive predictive value. This might explain the reporting of two causes of subfertility in the mailed questionnaire with confirmation of only one cause in the medical record. It is even possible that women reported causes of subfertility that were successfully treated in the past (e.g., tubal surgery), after which they received in vitro fertilization treatment (because of male subfertility) in the clinic participating in our study, to which the previously diagnosed cause of subfertility was not communicated. However, it seems unlikely to us that this has had more than a minor influence on our assessment of the accuracy of recall.

Our study was restricted to a subfertile population that received in vitro fertilization treatment. Recall of the cause of subfertility might be different among in vitro fertilization-treated women as compared with other subfertile women, because undergoing in vitro fertilization can be considered an important life event. In particular, women treated in the early era of assisted reproductive technology may remember their cause of subfertility more accurately because they undertook more extensive action to be treated. In addition, women treated with in vitro fertilization, compared with those subfertile women not receiving this treatment, might be more aware of reproductive health problems and might have higher educational levels (although in the Netherlands three in vitro fertilization attempts are covered by health insurance). These characteristics limit the generalizability of our results to subfertile women in general.

In the initial OMEGA study cohort, a subgroup of subfertile women who did not receive any treatment with in vitro fertilization was included ($n = 6{,}588$). However, because of limited project funding, we were able to abstract the medical records of only a small sample of these women who had not received in vitro fertilization treatment, and information on the cause of subfertility was absent in many cases. For 475 women not treated with in vitro fertilization, we had information on the cause of subfertility from both the questionnaire and the medical records. In this group, the positive predictive value, sensitivity, and kappa value for subfertility attributed to tubal factors were 94 percent, 91 percent, and 0.75, respectively (data not shown). These percentages are in line with those presented in table 2 for the in vitro fertilization-treated women.

The response rate to our mailed risk factor questionnaire was 71.2 percent. A survey among nonresponders showed that nonresponders were less likely to have had a livebirth after in vitro fertilization. The major reasons for nonresponse were lack of time to fill out the questionnaire and negative emotions regarding in vitro fertilization treatment, often related to the fact that no ongoing pregnancy was achieved. We cannot exclude the possibility that these reasons for nonresponse might be related to the accuracy of self-reported subfertility diagnosis.

In conclusion, our results indicate that self-reports of tubal and male subfertility are highly accurate. However, for unexplained subfertility, the validity of recall was considered to be only moderate and, for subfertility causes related to hormonal and uterine factors, the accuracy was considered low. The accuracy of recall was unsatisfactory if two causes of subfertility played a role. When self-reports on subfertility diagnosis are used in epidemiologic studies, verification by medical records should be a consideration for specific causes and combinations of more than one cause.

ACKNOWLEDGMENTS

The authors owe special thanks to Dr. H. Klip, who coordinated the identification of the cohort.

The OMEGA Project Group includes the following persons: Dr. M. Kortman and Dr. E. R. Te Velde (University Medical Center, Utrecht); Dr. N. Macklon (Erasmus Medical Center, Rotterdam); Dr. C. A. M. Jansen (Diaconessenhuis, Voorburg); Dr. R. A. Leerentveld (Isala clinics, Zwolle); Dr. W. N. P. Willemsen (Academic Hospital, Nijmegen, St. Radboud); Dr. R. Schats (Academic Hospital, Free University, Amsterdam); Dr. N. Naaktgeboren and Dr. F. M. Helmerhorst (Leiden University Medical Center, Leiden); Dr. R. S. G. M. Bots (St. Elisabeth Hospital, Tilburg); Dr. A. H. M. Simons (Academic Hospital, Groningen); Dr. H. V. Hogerzeil (Academic Medical Center, Amsterdam); Dr. J. L. H. Evers (Academic Hospital, Maastricht); and Dr. P. A. van Dop (Catharina Hospital, Eindhoven), the Netherlands.

REFERENCES

1. Whittemore AS, Harris R, Itnyre J. Characteristics relating to ovarian cancer risk: collaborative analysis of 12 US case-control studies. II. Invasive epithelial ovarian cancers in white women. Collaborative Ovarian Cancer Group. Am J Epidemiol 1992;136:1184–203.
2. Venn A, Watson L, Bruinsma F, et al. Risk of cancer after use of fertility drugs with in-vitro fertilisation. Lancet 1999;354:1586–90.
3. Burkman RT, Tang MT, Malone KE, et al. Infertility drugs and the risk of breast cancer: findings from the National Institute of Child Health and Human Development Women's Contraceptive and Reproductive Experiences Study. Fertil Steril 2003;79:844–51.
4. Klip H, van Leeuwen FE, Schats R, et al. Risk of benign gynaecological diseases and hormonal disorders according to

responsiveness to ovarian stimulation in IVF: a follow-up study of 8714 women. Hum Reprod 2003;18:1951–8.
5. Klip H, Burger CW, de Kraker J, et al. Risk of cancer in the offspring of women who underwent ovarian stimulation for IVF. Hum Reprod 2001;16:2451–8.
6. Brinton LA, Melton LJ 3rd, Malkasian GD Jr, et al. Cancer risk after evaluation for infertility. Am J Epidemiol 1989;129: 712–22.
7. Braga C, Negri E, La Vecchia C, et al. Fertility treatment and risk of breast cancer. Hum Reprod 1996;11:300–3.
8. Gammon MD, Thompson WD. Infertility and breast cancer: a population-based case-control study. Am J Epidemiol 1990;132:708–16.
9. Moseson M, Koenig KL, Shore RE, et al. The influence of medical conditions associated with hormones on the risk of breast cancer. Int J Epidemiol 1993;22:1000–9.
10. Rodriguez C, Tatham LM, Calle EE, et al. Infertility and risk of fatal ovarian cancer in a prospective cohort of US women. Cancer Causes Control 1998;9:645–51.
11. Harlow SD, Linet MS. Agreement between questionnaire data and medical records. The evidence for accuracy of recall. Am J Epidemiol 1989;129:233–48.
12. Paganini-Hill A, Ross RK. Reliability of recall of drug usage and other health-related information. Am J Epidemiol 1982; 116:114–22.
13. Olson JE, Shu XO, Ross JA, et al. Medical record validation of maternally reported birth characteristics and pregnancy-related events: a report from the Children's Cancer Group. Am J Epidemiol 1997;145:58–67.
14. Bean JA, Leeper JD, Wallace RB, et al. Variations in the reporting of menstrual histories. Am J Epidemiol 1979;109: 181–5.
15. Tilley BC, Barnes AB, Bergstralh E, et al. A comparison of pregnancy history recall and medical records. Implications for retrospective studies. Am J Epidemiol 1985;121:269–81.
16. Must A, Phillips SM, Naumova EN, et al. Recall of early menstrual history and menarcheal body size: after 30 years, how well do women remember? Am J Epidemiol 2002;155: 672–9.
17. de Boer EJ, Den Tonkelaar I, Te Velde ER, et al. Low number of retrieved oocytes at in vitro fertilisation treatment is predictive of early menopause. Fertil Steril 2002;77:978–85.
18. de Boer EJ, Den Tonkelaar I, Te Velde ER, et al. Increased risk of early menopausal transition and natural menopause after poor response at first IVF treatment. Hum Reprod 2003;18: 1544–52.
19. Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York, NY: John Wiley & Sons, Inc, 1981:217–19.
20. Speroff L, Glass RH, Kase NG. Female infertility. In: Mitchel C, Reter R, Stewart J, et al, eds. Clinical gynecologic endocrinology and infertility. Philadelphia, PA: Lippincott Williams & Wilkins,1999:1013–42.
21. Dick ML, Bain CJ, Purdie DM, et al. Self-reported difficulty in conceiving as a measure of infertility. Hum Reprod 2003;18: 2711–17.
22. Rich-Edwards JW, Goldman MB, Willett WC, et al. Adolescent body mass index and infertility caused by ovulatory disorder. Am J Obstet Gynecol 1994;171:171–7.