Improvement of accuracy of clinical reports—the case of IVF cycle rank

Yaakov Rosenfeld · Avshalom Strulov

Received: 26 June 2008 / Accepted: 14 January 2009 / Published online: 21 February 2009
© Springer Science + Business Media, LLC 2009

Abstract

Purpose Accurate clinical reports of In Vitro Fertilization Cycle Rank (IVF-CR) are important because of their predictive value of treatment success—clinical pregnancy. Our district currently provides more than a hundred IVF pick-up cycles per month. The study objective was to evaluate and improve the accuracy of IVF units' reports of IVF-CR.

Methods Reports of IVF-CR from local IVF units were correlated with administrative data. Couples who were having IVF treatments during the 3-month study period were included in the study. Data were collected before and after an intervention program.

Results During the study period accurate IVF-CR reporting rate improved from 27% (29 out of 108) to 83% (91 out of 110).

Conclusions Accuracy of clinical reports of IVF-CR from local units significantly improved after the intervention. It is recommended that national regulatory bodies should issue compulsory guidelines for the recording and reporting of IVF-CR.

Keywords Advanced reproductive techniques · Cost containment · Quality improvement · In Vitro Fertilization · Reliability of clinical reports

Capsule Lack of explicit guidelines about how IVF cycle should be counted is associated with inaccurate and unreliable reports on this parameter from local IVF unit.

Y. Rosenfeld (✉) · A. Strulov
School of Public Health, Haifa University,
Haifa, Israel
e-mail: r_yaakov@netvision.net.il

Y. Rosenfeld
District of Haifa and Western Galilee, General Health Services,
Haifa, Israel

Introduction

The right to reproduction is frequently considered as a natural right. Indeed, in most cases this right is naturally attained. In some cases, however, about 10% of the couples, it is not. In these cases, health authorities are called upon to intervene. While these interventions serve a supreme human cause, they also carry a heavy toll on the payer, be it private or public.

Public health care systems in developed countries found a variety of ways to deal with the escalating costs of Advanced Reproductive Technologies (ART's). While some do not relate to the problem of infertility as a "disease" and allow the interested couples to deal with costs issues themselves, others allow a limited number of In Vitro Fertilizations (IVF) cycles to be carried out on the expense of public funds (through a national health care system, or insurance).

In Israel, public health care is provided by four Health Maintenance Organizations (HMO's). These HMO's are funded by the state according to the number of enrollees they each register. Each HMO is then responsible to the care of its patients, either by its own employees, in its own facilities, or by purchasing services from secondary contractors. General Health Services is the largest HMO, with about 3.8 million members across the country. It is divided to eight regional operative units, called "districts." Israeli citizens are entitled to a uniform set of services (health basket), as dictated by the National Health Insurance Law. The basket includes IVF treatments for the first and second child. IVF units are generally reimbursed "per piece," i.e. per treatment cycle, according to the governmental tariff, with one exception. One unit is reimbursed "per success," i.e. per clinical pregnancy, at 12 weeks. Number of cycles is not limited. Rather, treatment is limited

by the woman's age (45 for a regular IVF, 52 for Egg Donation) [1]. Furthermore, treatment should stop after four consecutive cycles with either no ovum pick up, no fertilization, or no embryo transfer. With such flexible "constraints," it is not surprising that the state of Israel is found at the top of international comparisons with the highest per capita cycle number, high above the average of all the other industrialized countries.

One caveat still remains: how do we define a cycle? As obvious as the answer to this question may seem, it appears that no clear definition of how IVF cycles should be counted exists in the Israeli health care system.

The importance of accurate and reliable counting of IVF cycles is twofold: On the individual level, higher IVF cycle rank (IVF-CR) is considered by many to be associated with a lower success rate [2]. It is important therefore to know the actual IVF-CR beforehand, in order to present to the couple the realistic chances of having a baby. On the national level, it is important, in order to be able to delineate realistic policies that will direct public funds to couples who have reasonable chances to benefit from the treatment.

Under the current circumstances, of no explicit national guidelines of how IVF cycles should be counted, it was our hypothesis that IVF-CR, as reported by IVF units that provide this service to the District of Haifa and Western Galilee of General Health Services in Israel, is going to be neither accurate, nor reliable.

Objective

The primary purpose of this study was to evaluate the accuracy and validity of, IVF cycle rank as reported by attending physicians, from local IVF units, who provide this service to the enrollees from the District of Haifa and Western Galilee of the General Health Services. Should the results of this stage be non-satisfactory, stage II would be aimed to disclose underlying causes of these results, and phase III—to re-evaluate these parameters after an intervention program.

Materials and methods

A local committee, on the Districts level, has been operating in the District of Haifa and Western Galilee since January 1997. All the data pertinent to a certain couple are collected in a single file, throughout the period of advanced reproductive treatments. The members of the committee are two gynecologists, with expertise in fertility management, and one specialist in family medicine. The orientation of the committee is mainly clinical, rather than economical.

The study included three phases. Altogether, three non-consecutive months were studied, each one with a different focus, based on the results of the previous one.

Phase I. A sample cohort of one full month was evaluated by comparing cycle rank as reported by the IVF unit, with the data collected by the HMO.

The data collected and the comparisons were divided into a table testing its reliability from the two sources. Statistical analysis employed two Chi Square tests: one for "too high" reported cycle rank, and the other for "too low" reported cycle rank, as compared to the "exact" reported cycle rank.

Phase II. A second-month sample of current applications was then evaluated on an individual basis, in order to establish potential causes for discrepancies between reported and actual cycle rank. In light of the initial results, a feedback tool was developed, to inform local IVF units about correct IVF-CR in cases where reports were inaccurate.

Phase III. Finally, data on 110 consecutive applications from a third month were again collected during October–November 2008, in order to evaluate the impact of the feed-back mechanism. IVF-CR's as reported by local IVF units were again correlated with the data of the committee.

Results

Phase I. During the period of data collection, 108 IVF cycle treatments were reported. IVF-CR was reported for 69 patients (64%). For 39 patients (36%) data were missing. Accurate IVF-CR was reported for 29 patients (27%). For 30 patients (43.5%), reported IVF-CR was higher than the actual number, and for 10 (10%), it was lower. Chi Square test for "too high" reported cycle rank was highly significant, alpha 0.00000 with a power of 0.90. Chi Square test for "too low" reported cycle rank was also significant, alpha 0.042 with a power of 0.90.

Detailed differences between reported and actual IVF-CR are presented in Table 1 and in Fig. 1.

Phase II. Detailed examination of individual records from the second cohort revealed the following explanations for discrepancies between reported and actual IVF-CR:

IVF-CR which was higher than the actual number appeared in reports from units whose database included

Table 1 Differences between reported and actual IVF-CR

Delta	Number	Percent	Total High \ Low	Total Percent
16	1	0.9		
9	2	1.9		
7	0	0.0		
8	1	0.9		
6	1	0.9	30	27.8
5	5	4.6		
4	2	1.9		
3	3	2.8		
2	1	0.9		
1	14	13.0		
0	29	26.9	29	26.9
-1	5	4.6		
-2	2	1.9		
-3	1	0.9	10	9.3
-4	0	0.0		
-5	1	0.9		
-6	1	0.9	39	36.1
Missing				
Data	39	36.1		
Total	108	100	108	100

Delta = Reported—Actual IVF cycle rank;

Reported = 4, Actual = 2 => delta = +2

Reported = 2, Actual = 4 => delta = -2

Number: Number of cases in our sample

Total High: total number of reported IVF-CR > actual IVF-CR = 30/108

Total Low: total number of reported IVF-CR < actual IVF-CR = 10/108

one list of treatments, with no distinction between cycles with Ovum-Pick-Ups (OPU's) and cycles with Frozen-Embryo-Transfers (FET's).

In cases where reported IVF-CR was lower than the reported one, several reasons were identified:

- Transfer to a new unit—disregarding cycle treatments in previous units.
- A new calendar year—disregarding cycle treatments in previous years.
- Starting a new count after a previous treatment resulted in a clinical pregnancy and a live birth.

Phase III. At the third stage of the study, significant improvement of IVF-CR accuracy of reporting was noted: In 91 cases out of 108 (83%) IVF-CR was accurate. There was only one case of “too high,” and 18 cases (16%) of reported IVF-CR that were “too low.” Of these, in five cases (4.5%) it was only “minus one” too low. In the rest 13 cases (12%) reported IVF-CR was “too low” with a greater than one difference from the actual number.

Discussion

The importance of accurate and reliable reporting system for the processes and outcomes of advanced reproductive technologies needs no elaborate discussion. While the use of these techniques brings endless joy to couples who need them, they bear a heavy toll on the health budgets on different levels. The main reason for that is the sharp rise in resources directed to cover the costs of new treatments such as Intra-Cytoplasmic-Sperm-Injection (ICSI), Embryo-Cryo-Preservation and Thawing, and Testicular-Sperm-Aspiration (TESA).

An extensive international survey of the health economics of IVF and ICSI [3] estimated utilization rate of IVF/ICSI in 48 countries to be 289 cycles per million per year. National numbers ranged from 126 in the USA, 180 and 190 in Brazil and Canada, 441 and 610 in the UK and France, to 829 and 899 in the Netherlands and Iceland. The highest rate was found in Israel—1,657, and the lowest in Kazakhstan—2. A wide range of costs was also found in this survey: average cost was estimated to be US\$3,518 (projection to 2002) in the 25 reporting countries outside the USA, where the cost was estimated to be US\$9,547. The costs (in 2002 US\$) were estimated to be \$4,532 in Canada, \$2,955 in the UK, and \$3,950 in Iceland. The most expensive was reported from Hong-Kong—\$6,361, and the cheapest from Pakistan—\$1,272. Data on costs were missing for 22 of the 48 reporting countries including Brazil, France and the Netherlands.

In Israel, Stern et al. [4] estimated the cost of an IVF cycle treatment to be US\$2,560, and the cost of achieving the desired result (“take-home baby”), taking into account the success rate at that time (15%), to be US\$19,267. The same team of investigators revisited the subject later [5], after the introduction of the micro-manipulation technology. The updated calculation estimated the cost of a cycle treatment with ICSI to be 61% higher than that of a standard IVF, with the cost of “take-home baby” 54% higher.

As appealing as these articles may be, they all share a common assumption: they rely on data reported by IVF units, and assume that these data are valid and accurate. This hypothesis, however, has never been tested. Extensive

Fig. 1 Differences between reported and actual IVF-CR

computerized literature search did not yield a single report of a study that explores this supposition. Furthermore, implicit uncertainty about the accuracy of the data supplied by IVF units was expressed in two cases: Jain et al [6], in their report about “Insurance coverage and outcomes of IVF,” in the Methods section, reported that “Despite the federal requirements to report success rates, 30 clinics [out of 390–7.7%] either failed to submit their data to the CDC or did not provide verification by the clinic’s medical director that the tabulated success rates were correct”.

In Israel, in 1989, the Ministry of Health appointed a National Committee for Collection and Registration of Advanced Reproductive Techniques Data. Since then, the Committee has published its reports on several occasions [7, 8, and 9] with the last publication referring to data of the years 1995–96 [10] on December 2000. Data collection by the Committee is based on questionnaires distributed to the IVF units throughout the country. The question of the reliability of the data reported to the Committee had been raised earlier [9], but data collection mode remained unchanged.

Testing the hidden assumption of the validity of data reported by IVF units seemed, therefore, justified and required. IVF cycle rank, which was considered a straight-forward parameter to record and report, was selected intentionally, rather than “sensitive” parameters, such as embryo quality or even follicular diameter, which are prone to inter-observer variation. Cycle rank, which is a counting task, seemed simple enough, yet important, to serve as a parameter of the validity of the clinical reports. The results of phase I of our study demonstrate that the IVF-CR, as reported by the caregivers, were valid and accurate in just a minority of the cases: In our study, clinical reports of IVF-CR were accurate in only 27% of the cases. For 30 patients (43.5%) it was higher than the actual number, and for 10 (10%), lower. In 39 cases (36%) IVF units were reluctant to relate to this basic information, in spite of specific requests. This figure of non-response rate is much higher than that of Jane’s study, probably due to differences in the source of authority: a customer, in our study, as opposed to a federal agency in theirs.

Phase II of the study demonstrated that while a number of specific reasons for inaccurate reports of IVF-CR were detected, a common root cause seems to be involved in all cases, that is, the lack of an updated and formal reporting system. While the stages of IVF treatment are well defined, mainly for accounting purposes, the question of how IVF cycles should be counted was left unanswered. This fact enables IVF units to artificially “increase” IVF cycle numbers (by counting “fresh” and “frozen” cycles together) or “decrease” this number (by starting a “new” count after an arbitrary event).

Phase III. At the third stage of the study, significant improvement of IVF-CR accuracy of reporting was noted. Accurate IVF-CR were now reported for 91 cases out of

108, (83%) compared to only 27% accurate IVF-CR at phase I. It seems that our concept of counting OPU cycles in one list, and FET’s cycles in a separate list was generally accepted. The one case of “too high” was reported from a remote hospital with a very low work load with us. The 18 cases (16%) that demonstrated “too low” reported IVF-CR seemed to share two mechanisms: in five cases (4.5%), reported IVF-CR was only minus one “too low”. This minor inaccuracy seemed to stem from the fact that in our files, at the application/pre-authorization stage, IVF-CR relates to next cycle, the one that is considered for approval, while reporting physicians probably related to the IVF-CR that were already performed. In the remaining 13 cases (12%), reported IVF-CR was “too low” with a greater than one difference from the actual number. All these cases had one common inspiring denominator: in all these cases, cycle treatments in previous units were ignored.

Generalization of the results of our study is limited: Our data covers only the activity in the public system. Should a couple choose to have their treatment entirely in the private system (for reasons of confidentiality, for example)—we would not know about it. In our experience, however, that this is the case in only a negligible proportion of the cases. Furthermore, should a case like this happen—it would be missing from both the nominator and the denominator, having small effect on the net result. Other cases which are not represented in our data are cases that are not covered by the insurer, such as treatment for the third child, or treatment of a woman over 45 years old. Again, as these treatments are not included, neither in our database, nor in the reports, this should not affect the net result.

It can be argued that our data represent the reality of only a narrow time frame, in a small geographic area. It should be noted, however, that our database was established on January 1997, more than a decade ago. The local IVF control system was adopted by the central management of General Health Services to its other seven districts throughout the country a couple of years later. It can, therefore, be assumed, with a high degree of confidence, that if any difference exists between the levels of completeness and accuracy of our data, and those of other districts, in other time frames, the latter would be inferior. In a clinical framework that enabled GJ Annas to name his article [11] “The shadowlands—secrets, lies and assisted reproduction,” our limited data may serve as an initial step to shed some light, to clear the fog, and lift the veil of uncertainty, in favor of a better understanding of what is really going on.

It is our opinion that a significant and straight-forward counting system can be achieved if OPU’s are counted in one list and FET’s in a separate list. Any “intention to treat” should be counted in order to include “dropped cycles.” The counting should begin with the first treatment, and end

with the last, regardless of whether the results of any specific treatment were satisfactory or unsatisfactory, whether a couple chose to continue the treatment in another unit or whether a new year begun. Any list that counts together pick-ups and transfers, fresh cycles and frozen-thawed cycles, self with donated gametes is doomed to create confusion and a hopelessly chaotic database. It is recommended that national regulatory agencies issue compulsory directions with regard to recording and reporting of IVF cycle rank, and establish a mechanism to verify that these reports are indeed accurate and reliable.

Rational process of decision making by policy makers is dependent on the existence of such a system.

References

1. Internal Memorandum, 19/99 pp. 7–8. Central management—general health services.
2. Padilla SL, Garcia JE. Effect of maternal age and number of in vitro fertilization procedures on pregnancy outcome. *Fertility and Sterility*. 1989;52(2):270–3. Aug.
3. Collins JA. An international survey of the health economics of IVF and ICSI. *Human Reproduction Update*. 2002;8(3):265–77.
4. Stern Z, Laufer N, Levy R, Ben-Shushan D, Mor-Yosef S. Cost analysis of In Vitro fertilization. *Israel J Med Sci*. 1995;31(8):492–6.
5. Ezra Y, Ben-Shushan D, Scenker J. The cost of micromanipulative assisted reproduction: new aspects to consider. *Israel J Obstet Gynecol*. 1998;9(2):99–103.
6. Jain T, Harlow BL, Hornstein MD. Insurance coverage and outcomes of In Vitro Fertilization. *N Eng J Med*. 2002;9:661–6. August, 374.
7. Lotan Y, Seidman DS. In vitro fertilization and embryo transfer in Israel: results from a national survey. *J Assist Reprod Genet*. 1992;9:419–21.
8. Seidman DS, Lotan Y. In vitro fertilization and embryo transfer in Israel: 1990 results from a national survey. *J Assist Reprod Genet*. 1994;11:1–4.
9. Seidman DS, Lotan Y. Assisted reproductive technology in Israel: results of the national registry up to 1994. *Israel J Obstet Gynecol*. 1998;9:18–24.
10. Insler V, Gonen O, Levran D, Lotan Y, Fish B, Potashnik G, Kogosovsky A, Ron-El R. Assisted reproductive technologies reported in the national registry of Israel 1995–1996. *Harefuah*. 2000;139(11–12):421–4. Dec.
11. Annas GJ. The shadowlands—secrets, lies and assisted reproduction. *N Eng J Med* September. 1998;339:935–9.