Reproductive BioMedicine Online (2010) **20**, 533–542

www.sciencedirect.com
www.rbmonline.com

ARTICLE

Variation in guideline adherence in intrauterine insemination care

Esther C Haagen [a,*], Willianne LDM Nelen [a], Richard PTM Grol [b], Didi DM Braat [a], Rosella PMG Hermens [b], Jan AM Kremer [a]

[a] *Department of Obstetrics and Gynaecology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands;*
[b] *Scientific Institute for Quality of Healthcare (IQ healthcare), Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands*
[*] Corresponding author. *E-mail address:* e.haagen@obgyn.umcn.nl (EC Haagen).

Esther Haagen is a third-year gynaecologist-in-training at the Department of Obstetrics and Gynaecology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. She is currently undertaking a PhD at the Radboud University Nijmegen Medical Centre within a research group interested in clinical practice guidelines, performance indicators and implementation strategies to improve infertility care. Her research is focused on intrauterine insemination care.

Abstract Health-care delivery according to clinical practice guidelines is thought to be critical in achieving optimal outcomes. This study aimed to assess the extent to which practice performance in intrauterine insemination (IUI) care is consistent with guideline recommendations and to evaluate the association between guideline adherence and outcome of IUI care. In a retrospective cohort study, 1100 infertile couples who underwent IUI treatment at 10 Dutch hospitals were asked to grant access to their medical record for assessment of guideline adherence using 25 systematically developed guideline-based performance indicators. A total of 558 couples who started 2334 IUI cycles participated. Guideline adherence regarding 20 process and five structure aspects of IUI care was often substandard and varied considerably between hospitals. Out of 10 possible associations investigated, guideline adherence regarding sperm quality and guideline adherence regarding the total number of IUI cycles were associated with improved ongoing pregnancy rates after IUI. Thus, guideline adherence in IUI care is far from optimal and varies substantially between hospitals. As associations between guideline adherence and ongoing pregnancy after IUI were mainly non-significant, further research is needed to evaluate associations between guideline adherence and other outcomes of IUI care besides ongoing pregnancy, such as patient safety and cost effectiveness.

© 2009, Reproductive Healthcare Ltd. Published by Elsevier Ltd. All rights reserved.

KEYWORDS: guidelines, indicators, infertility, intrauterine insemination, outcome, performance

Introduction

Infertility is common. As one in six couples fails to conceive within 1 year of regular unprotected sexual intercourse, approximately 72.4 million couples around the world are currently infertile (Boivin et al., 2007; Taylor, 2003). Roughly half of these couples seek medical care for their fertility problems (Boivin et al., 2007; Nachtigall, 2006).

1472-6483/$ - see front matter © 2009, Reproductive Healthcare Ltd. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.rbmo.2009.12.016

The clinical investigation and treatment of infertility has considerable physical and psychosocial implications for couples (Cousineau and Domar, 2007; Greil, 1997; Hughes, 2003; Verhaak et al., 2005). In addition, fertility treatment often carries a risk of multiple pregnancy which is associated with substantially higher rates of maternal and perinatal morbidity and mortality compared with singletons (Dickey et al., 2005; Fauser et al., 2005). The assessment and treatment of infertility is also expensive and the increased rates of adverse pregnancy outcomes add further to the resources used in health care (Goverde et al., 2000; Lukassen et al., 2004).

It is widely recognized that the quality of health care is often inadequate (Grol and Grimshaw, 2003; McGlynn et al., 2003). As infertility is a major health issue worldwide and its management is associated with health risks and high costs, there is great concern about the impact of substandard infertility care on health and health-care resources (Balen and Rutherford, 2007; Rawal et al., 2008). This prompted several professional societies in the fields of obstetrics and gynaecology, and reproductive medicine as well as governmental agencies to develop clinical practice guidelines for infertility care. These guidelines are tools to help physicians and patients to make better decisions about clinically effective, safe and cost-effective care and reduce inappropriate practice variation in infertility care. Essentially, these guidelines describe 'the best thing to do' and aim to improve the outcome of infertility care.

One of the most commonly used fertility treatments is homologous intrauterine insemination (IUI) (Nyboe Andersen et al., 2009). The overall success rate of IUI remains controversial. On average, reported clinical pregnancy rates are only 5—13% per IUI cycle (Bensdorp et al., 2007; Goverde et al., 2000; Guzick et al., 1999; Nyboe Andersen et al., 2009; Steures et al., 2004, 2007; The ESHRE Capri Workshop Group, 2009; Tummon et al., 1997; Verhulst et al., 2006). Poor guideline adherence and a high degree of inappropriate practice variation in IUI care may be responsible for these low pregnancy rates. However, up until now, there has been limited published evidence about guideline adherence in IUI care and the link between guideline adherence and outcome of IUI care. Therefore, the objective of this study was to assess the extent to which practice performance in IUI care is consistent with guideline recommendations and to evaluate the association between guideline adherence and outcome of IUI care.

Materials and methods

Study design and population

This retrospective cohort study at 10 Dutch hospitals was conducted using medical record and questionnaire data. The group of participating hospitals was made up of one large academic hospital providing tertiary care and nine medium-sized public hospitals providing secondary care. Five clinics were also teaching hospitals. Patients eligible to participate in the study were defined as infertile couples who underwent IUI treatment at these 10 hospitals during an inclusion period of 28 months, from April 2000 to August 2002. Infertile couples who had undergone ovulation induction therapy for ovulatory disorders or IUI treatment with donated spermatozoa were not eligible to participate. Once the relevant institutional review boards gave their approval, the databases of the fertility laboratories of the participating hospitals were used to make a list of eligible couples per hospital. From each list, a random sample of couples in proportion to the size of the hospital's IUI programme (50—150 couples per hospital) was selected by computer for participation in the study. The total sample consisted of 1100 infertile couples. Each couple was requested in an informative letter to sign and return a consent form to grant access to their medical record. Non-respondents were sent a reminder after 4 weeks.

Indicator development

Reliable assessment of the extent to which practice performance in health care is consistent with guideline recommendations requires a set of valid guideline-based performance indicators (Grol et al., 2002; Nelen et al., 2007). Therefore, a set of guideline-based performance indicators for IUI care was systematically developed for use in three steps: (i) preselection of recommendations; (ii) expert consensus procedure; and (iii) transcription and classification of final set of indicators (Campbell et al., 2003; Mainz, 2003; Rubin et al., 2001; Schouten et al., 2005).

First, three authors independently preselected all recommendations from the four existing IUI guidelines in Europe (Danish Fertility Society (DFS), 2007; Dutch Society of Obstetrics and Gynaecology (NVOG), 2000; National Agency for Accreditation and Evaluation in Healthcare (ANAES), 1996; National Institute for Health and Clinical Excellence (NICE), 2004). These IUI guidelines from Denmark, the Netherlands, France, and England and Wales are evidence-based and of moderate to high methodological quality according to appraisal with the internationally validated AGREE Instrument (Haagen et al., 2006). Discrepancies in the recommendations of the four IUI guidelines were assessed against the best available evidence to ensure selection of recommendations supported by the highest level of evidence. This resulted in a list of 33 recommendations on IUI care.

Second, an expert consensus procedure according to the RAND-modified Delphi method was conducted to explore consensus among a group of opinion leaders about the relevance of the preselected recommendations (Campbell et al., 2000, 2003; Cantrill et al., 1998; Rubin et al., 2001). In the two-round consensus procedure, a panel of 13 experts in the field of IUI and quality-of-care research rated the relevance of the 33 preselected recommendations to patients' health benefit and cost effectiveness on a five-point Likert scale (1 = strongly disagree; 5 = strongly agree). Moreover, the experts were asked to provide a top-five ranking of recommendations they considered most important. The experts also had the opportunity to comment on each proposed recommendation and suggest additional recommendations for evaluation. In the first round, 10 recommendations with at least 75% of the ratings on both criteria in the categories 'agree' and 'strongly agree' and a top-five ranking were selected. In the second round, the experts re-evaluated the remaining 23 recommendations and eight newly suggested recommendations. This re-evaluation resulted in the selection of another 15 recommendations with at least 75% of the ratings on one criterion in the categories 'agree' and 'strongly agree'.

Third, the new selection of 25 key recommendations on IUI care was transcribed into indicators. This resulted in the final set of 25 guideline-based performance indicators for IUI care. Of these, 20 indicators were related to process aspects of IUI care, subdivided into 10 main categories: screening for tubal occlusion ($n = 1$), sperm quality ($n = 1$), indications for (un)stimulated IUI ($n = 3$), total number of IUI cycles ($n = 2$), monitoring in (un)stimulated IUI ($n = 3$), timing in (un)stimulated IUI ($n = 2$), dose of gonadotrophins in stimulated IUI ($n = 3$), timing of human chorionic gonadotrophin (HCG) administration in stimulated IUI ($n = 2$), dose of HCG in stimulated IUI ($n = 1$), and cancellation criteria in stimulated IUI ($n = 2$) (**Table 1**). The other five indicators referred to structure aspects of IUI care: practice facilities for IUI ($n = 5$) (**Table 2**).

Data collection

Data about process aspects of IUI care (20 process indicators described in **Table 1**), clinical outcomes and patient characteristics, such as female age, type and duration of infertility and diagnosis, were abstracted from medical records by two trained research assistants using standardized audit forms. To assess the reliability of data collection, both research assistants independently re-abstracted a randomly selected sample of 24 medical records from two different hospitals. The extent of agreement regarding the collected data, corrected for chance, was calculated using the kappa statistic. The average reliability of data collection was very good ($\kappa = 0.96$; range 0.83–1.0).

Data about structure aspects of IUI care (five structure indicators described in **Table 2**) were collected by means of a short questionnaire distributed to one gynaecologist in each participating hospital.

Statistical analysis

To assess the extent to which practice performance in IUI care is consistent with guideline recommendations, the performance scores of the 25 indicators for IUI care were calculated.

The performance score of each process indicator was calculated in two steps. First, the number of IUI cycles of a couple in which practice performance was consistent with a guideline recommendation was divided by the total number of IUI cycles of the couple to which the recommendation applied. Second, these proportions of individual infertile couples were added together and divided by the number of infertile couples to whom the guideline recommendation applied. If IUI treatment resulted in a pregnancy, only IUI cycles prior to this pregnancy were used for analysis, excluding possible consecutive IUI cycles within the study period.

The performance score of each structure indicator was calculated by dividing the number of hospitals in which practice performance was consistent with a guideline recommendation by the number of hospitals to which the recommendation applied. Performance scores ranged from 0 to 100%.

To evaluate the association between guideline adherence and outcome of IUI care, survival analysis based on the Cox proportional hazards model was used. The Cox proportional hazards model makes it possible to estimate the effect of explanatory variables on time to events, taking incomplete follow-up into account. Ongoing pregnancy, defined as a clinical pregnancy confirmed by the presence of fetal heart rate beyond 16 weeks' gestation, was considered the main outcome and selected as dichotomous censoring variable (ongoing pregnancy or no ongoing pregnancy). The 20 process indicators for IUI care were eligible explanatory variables. However, the indicators within each of the 10 main categories of process indicators were aggregated to reduce the number of explanatory variables from 20 to 10. Accordingly, 10 aggregated process indicators (**Table 1**) were selected as dichotomous explanatory variables (performance score $<90\%$ or $\geq 90\%$). Four patient and four hospital characteristics were taken into account as cofactors: female age (years), type of infertility (primary or secondary), duration of infertility (months), diagnosis (male factor fertility problems, unexplained fertility problems or cervical mucus hostility), hospital size (<700 or ≥ 700 beds), teaching hospital (teaching or non-teaching), IVF facilities (IVF or no IVF), and number of physicians involved in the IUI programme (≤ 10 or >10).

Two separate survival models were designed. One model for the total cohort and a second model for a subgroup who underwent stimulated IUI treatment since several process indicators were solely related to stimulated IUI, as opposed to unstimulated IUI. $P < 0.05$ was considered statistically significant. Analyses were performed using SAS software (SAS 8.2 for Windows; SAS Institute, Cary, North Carolina, USA).

Results

Study population

Figure 1 presents the recruitment of eligible infertile couples for participation in the study. A total of 765 infertile couples was willing to participate. The study excluded 184 couples because they had undergone ovulation induction therapy for ovulatory disorders or IUI treatment with donated spermatozoa. Another 23 couples were excluded because there was no access to their medical records. As a result, 558 infertile couples who started a total of 2334 IUI cycles were eligible for study.

Table 3 shows the patient characteristics. Median female age at the first IUI cycle was 32 years (range 23–46). Most couples (49%) were diagnosed with unexplained infertility. The median number of started IUI cycles was three (range 1–15). The patient characteristics female age, type of infertility (primary or secondary), duration of infertility and diagnosis (male factor fertility problems, unexplained fertility problems or cervical mucus hostility) did not differ significantly between hospitals.

Follow-up until the pre-specified outcome, either pregnancy or discontinuation of unsuccessful IUI treatment, was complete for 415 couples who started a total of 1803 IUI cycles. A total of 131 of these couples achieved an ongoing pregnancy after IUI treatment, giving an ongoing pregnancy rate per followed-up couple of 31.6% (range across 10 hospitals, 17–49%). Among the 131 IUI-related ongoing pregnancies were 13 multiples (9.9%; range across

Table 1 Performance scores of 20 process indicators for intrauterine insemination care at 10 hospitals.

Process indicator	*No. of infertile couples to whom indicator applies*[a]	*Performance score (%)*[b]	*Range across 10 hospitals (%)*
Screening for tubal occlusion			
Before starting IUI, screening for tubal occlusion should be performed	558	75	66–93
Sperm quality			
More than one million motile spermatozoa should be available for IUI after sperm preparation	553	74	0–88
Indications for (un)stimulated IUI			
Couples with male factor fertility problems should be offered unstimulated IUI	40	59	0–100
Couples with unexplained fertility problems should be offered stimulated IUI	208	94	70–100
Couples with cervical mucus hostility should be offered unstimulated IUI	51	38	0–100
Total number of IUI cycles			
Couples with male factor fertility problems should be offered up to six IUI cycles	78	26	0–43
Couples with unexplained fertility problems should be offered up to six IUI cycles	79	46	0–73
Monitoring in (un)stimulated IUI			
Patients undergoing stimulated IUI should be monitored by transvaginal ultrasonography	450	98	87–100
Patients undergoing unstimulated IUI should be monitored by LH measurements	173	53	0–100
LH measurements should be performed twice daily	110	25	0–66
Timing in (un)stimulated IUI			
IUI should be performed 38–42 h after administration of HCG	429	19	3–68
IUI should be performed 20–30 h after detection of spontaneous LH surge	241	9	0–42
Dose of gonadotrophins in stimulated IUI			
The dose of gonadotrophins in the first cycle of stimulated IUI should be 75 IU per day	254	30	0–63
The dose of gonadotrophins should be adapted if ovarian stimulation does not result in two or three follicles larger than 16 mm	277	33	13–47
If the dose of gonadotrophins is raised to achieve multifollicular growth, this should be done with 37.5 IU per day per cycle	23	20	0–79
Timing of HCG administration in stimulated IUI			
Patients undergoing stimulated IUI with gonadotrophins should be administered HCG when dominant follicular diameter reaches 18 mm	247	47	26–73
Patients undergoing stimulated IUI with oral anti-oestrogens should be administered HCG when dominant follicle diameter reaches 20 mm	200	48	0–75

Table 1 (*continued*)

Process indicator	*No. of infertile couples to whom indicator applies*[a]	*Performance score (%)*[b]	*Range across 10 hospitals (%)*
Dose of HCG in stimulated IUI			
HCG should be administered at a dose of 5000 IU	403	21	0—100
Cancellation criteria in stimulated IUI			
An IUI cycle should be cancelled if ovarian ultrasound reveals five or more follicles >12 mm or three or more follicles >16 mm	42	71	0—100
Advice to withhold from unprotected sexual intercourse should be given if ovarian ultrasound reveals five or more follicles >12 mm or three or more follicles >16 mm	30	53	0—100

HCG = human chorionic gonadotrophin; IUI = intrauterine insemination; stimulated IUI = IUI after ovarian stimulation using gonadotrophins or oral anti-oestrogens; unstimulated IUI = IUI in natural cycles; LH = luteinizing hormone.
[a]The maximum number of infertile couples to whom a process indicator can apply is 558.
[b]Each performance score was calculated in two steps: (i) the number of IUI cycles of a couple in which practice performance was consistent with a guideline recommendation was divided by the total number of IUI cycles of the couple to which the recommendation applied; and (ii) the proportions of individual infertile couples were added together and divided by the number of infertile couples to whom the guideline recommendation applied.

Table 2 Performance scores of five structure indicators for intrauterine insemination care at 10 hospitals.

Structure indicator	*Performance score (%)*
Practice facilities for IUI	
Semen volume for IUI should be 0.2—0.5 ml after sperm preparation	90
IUI laboratory should be accredited	60
Ovarian ultrasound monitoring to measure follicular size and number should be possible every day	90
IUI should be possible at least 5 days per week	100
IUI treatment results should be evaluated yearly	50

10 hospitals, 0—29%), including 12 twins and one triplet. All multiple pregnancies resulted from IUI combined with ovarian stimulation using gonadotrophins (77%) or oral anti-oestrogens (23%).

Guideline adherence regarding process aspects of IUI care

Guideline adherence regarding the 20 process aspects of IUI care varied substantially (**Table 1**). Adherence to recommended processes ranged from 9% for timing in unstimulated IUI to 98% for monitoring by transvaginal ultrasonography in stimulated IUI. Several process aspects, including the total number of IUI cycles, timing in (un)stimulated IUI, the dose of gonadotrophins, timing of HCG administration and the dose of HCG in stimulated IUI, had low rates of adherence (performance scores 9—48%).

Adherence to recommended processes also differed largely between the 10 participating hospitals. Differences between hospitals were particularly striking with regard to indications for (un)stimulated IUI and cancellation criteria in stimulated IUI, with performance scores ranging from 0 to 100%. Typical differences existed between hospitals in guideline adherence regarding the dose of HCG in stimulated IUI. While two hospitals always administered HCG at the recommended dose of 5000 IU (performance score 100%), the other eight hospitals always or predominantly administered HCG at a dose of 10.000 IU (performance score 0—25%).

Guideline adherence regarding structure aspects of IUI care

Guideline adherence regarding the five structure aspects of IUI care varied as well (**Table 2**). Appropriate practice facilities for IUI provided infertile couples in the majority

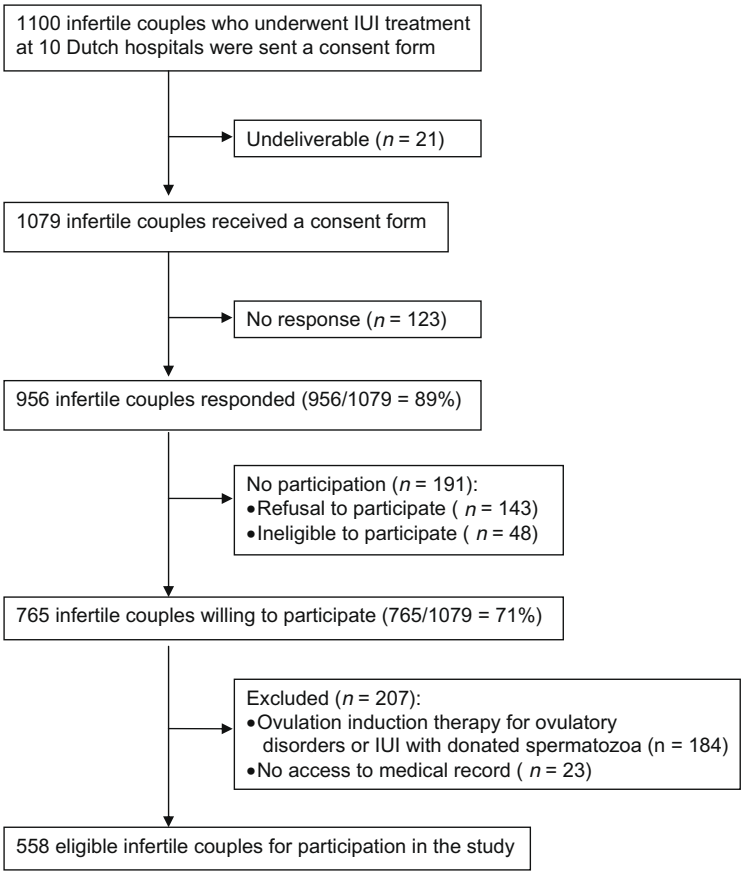

Figure 1 Recruitment of eligible infertile couples for participation in the study.

Table 3 Patient characteristics (n = 558).

Characteristic	*Value (range)*
Median female age (years)	32 (23–46)
Type of infertility (%)	
Primary	66
Secondary	34
Median duration of infertility (months)[a]	26 (0.3–112)
Diagnosis (%)	
Male factor fertility problems	42
Unexplained fertility problems	49
Cervical mucus hostility	9
Median number of started IUI cycles	3 (1–15)
Number of IUI-related ongoing pregnancies[b]	
Singletons	118
Multiples	13

IUI = intrauterine insemination.
[a]Duration of infertility was defined as the period between the start of regular unprotected sexual intercourse and the first IUI cycle.
[b]Ongoing pregnancy was defined as a clinical pregnancy of more than 16 weeks' gestation.

of hospitals the opportunity to receive recommended care, including the most favourable semen volume for IUI (performance score 90%), ovarian ultrasound monitoring on every day of the week (performance score 90%) and IUI on at least 5 days per week (performance score 100%). However, structure aspects such as the accreditation status of IUI

Table 4 Predictors of an ongoing pregnancy after intrauterine insemination (IUI) in the total cohort ($n = 558$).

Determinant	β	P-*value*	*Hazard ratio (95% CI)*	*Predictors of ongoing pregnancy*[a]
Process indicator				
Sperm quality[b]	1.715	<0.0001	5.56 (3.74–8.25)	More than one million motile spermatozoa
Total number of IUI cycles[c]	−1.493	<0.0001	0.23 (0.11–0.46)	Not more than six IUI cycles
Patient characteristic				
Type of infertility[d]	−0.386	0.0376	0.68 (0.47–0.98)	Secondary infertility
Duration of infertility[e]	−0.014	0.0381	0.98 (0.97–0.99)	Shorter duration of infertility

[a]Ongoing pregnancy was defined as a clinical pregnancy of more than 16 weeks' gestation.
[b]Process indicator regarding sperm quality: 'More than one million motile spermatozoa should be available for IUI after sperm preparation'.
[c]Process indicator regarding the total number of IUI cycles: 'Couples should be offered up to six IUI cycles'.
[d]Infertility can be primary or secondary.
[e]Duration of infertility was defined as the period between the start of regular unprotected sexual intercourse and the first IUI cycle.

laboratories and yearly evaluation of IUI treatment results had lower rates of adherence (performance scores respectively 60 and 50%).

Guideline adherence and ongoing pregnancy after IUI

Survival analysis revealed three statistically significant associations between ≥90% guideline adherence and ongoing pregnancy after IUI, out of 10 possible associations investigated (**Tables 4 and 5**). Guideline adherence regarding sperm quality had the strongest and most consistent association with improved ongoing pregnancy rates after IUI. Adherence to the recommendation that more than one million motile spermatozoa should be available for IUI after sperm preparation increased the chance of an ongoing pregnancy more than five to six times (hazard ratio 5.56 (3.74–8.25) in total cohort; hazard ratio 6.05 (3.69–9.93) in subgroup who underwent stimulated IUI, both $P < 0.0001$). Guideline adherence regarding the total number of IUI cycles was also associated with improved ongoing pregnancy rates after IUI. Indeed, adherence to the recommendation that couples should be offered up to six IUI cycles increased the chance of an ongoing pregnancy and the likelihood of an ongoing pregnancy was considerably reduced if infertile couples continued IUI treatment beyond the recommended number of six cycles (hazard ratio 0.23 (0.11–0.46), $P < 0.0001$ in total cohort; hazard ratio 0.21 (0.10–0.47),

Table 5 Predictors of an ongoing pregnancy after intrauterine insemination (IUI) in a subgroup who underwent stimulated IUI ($n = 400$).

Determinant	β	P-*value*	*Hazard ratio (95% CI)*	*Predictors of ongoing pregnancy*[a]
Process indicator				
Sperm quality[b]	1.800	<0.0001	6.05 (3.69–9.93)	More than one million motile spermatozoa
Total number of IUI cycles[c]	−1.546	0.0001	0.21 (0.10–0.47)	Not more than six IUI cycles
Dose of HCG[d]	−0.705	0.0411	0.49 (0.25–0.97)	More than 5000 IU HCG
Patient characteristic				
Duration of infertility[e]	−0.030	0.0008	0.97 (0.95–0.99)	Shorter duration of infertility

[a]Ongoing pregnancy was defined as a clinical pregnancy of more than 16 weeks' gestation.
[b]Process indicator regarding sperm quality: 'More than one million motile spermatozoa should be available for IUI after sperm preparation'.
[c]Process indicator regarding the total number of IUI cycles: 'Couples should be offered up to six IUI cycles'.
[d]Process indicator regarding the dose of HCG in stimulated IUI: 'HCG should be administered at a dose of 5000 IU'.
[e]Duration of infertility was defined as the period between the start of regular unprotected sexual intercourse and the first IUI cycle.

$P = 0.0001$ in subgroup who underwent stimulated IUI). In contrast, adherence to the recommended dose of 5000 IU HCG in stimulated IUI decreased the chance of an ongoing pregnancy by half (hazard ratio 0.49 (0.25–0.97), $P < 0.0441$ in subgroup who underwent stimulated IUI).

Survival analysis also showed that two patient characteristics were significantly related to ongoing pregnancy after IUI (**Tables 4 and 5**). The likelihood of an ongoing pregnancy was much lower if couples had never conceived (hazard ratio 0.68 (0.47–0.98), $P = 0.0376$ in total cohort) or had a longer duration of infertility (hazard ratio 0.98 (0.97–0.99), $P = 0.0381$ in total cohort; hazard ratio 0.97 (0.95–0.99), $P = 0.0008$ in subgroup who underwent stimulated IUI).

Discussion

Although guideline adherence appears adequate for some process and structure aspects of IUI care, on the whole, guideline adherence in IUI care is far from optimal. Also striking is the large variability in guideline adherence between different hospitals. Associations between ≥90% guideline adherence and ongoing pregnancy after IUI were mainly non-significant.

Strengths and limitations of the study

The strength of the study is the rigorous assessment of the extent to which practice performance in IUI care is consistent with guideline recommendations, using a systematically developed set of valid guideline-based performance indicators for IUI care. The large sample of 558 infertile couples from 10 different hospitals who started a total of 2334 IUI cycles contributes to the validity of these research results.

The study also has some limitations. First, the influence of selection bias remains uncertain (Grimes and Schulz, 2002). Twenty-five per cent of the selected couples refused to participate in the study. More than half of these couples reported the reason for refusal to participate, mainly lack of time or interest. As the overall participation rate was adequate, it is suspected that selection bias did not significantly distort the results of the study.

Second, the study relied primarily on chart abstraction to collect data on practice performance in IUI care. This was an elaborate exercise and may also have introduced bias as previous research showed that chart abstraction somewhat underestimates the quality of actual practice performance (Luck et al., 2000). However, in the near future, the use of electronic medical records will increase, presenting an excellent opportunity to incorporate performance indicators in electronic medical records as mandatory items to fill out. This will almost certainly facilitate the assessment, registry, monitoring and comparison of practice performance on local, national and international level (Nelen et al., 2007; Nyboe Andersen et al., 2009; Steures et al., 2007).

Third, incomplete follow-up could also be considered a limitation of the study. Ideally, all infertile couples enrolled in the study are followed until the pre-specified outcome, either pregnancy or discontinuation of unsuccessful IUI treatment, is observed. However, some couples had not yet reached this outcome by the time the study ended. Therefore, survival analysis based on the Cox proportional hazards model, which takes this incomplete follow-up properly into account, was used.

Fourth, to evaluate the association between guideline adherence and ongoing pregnancy after IUI, the study performed survival analysis in which guideline adherence was defined as adherence to a guideline recommendation in ≥90% of IUI cycles per couple. This cut-off point of ≥90% may seem arbitrary. However, as survival analysis was performed using other cut-off points, such as ≥75%, ≥80% and ≥85%, the survival model did not change, just the estimates of the effect to a limited extent.

Meaning of the study

Published literature is full of examples of patients receiving inappropriate care according to current scientific evidence and professional insight laid down in guidance (Campbell et al., 2005; Grol and Grimshaw, 2003; McGlynn et al., 2003; Sheldon et al., 2004). This study established that practice performance in IUI care is for the most part not consistent with guideline recommendations. There is also a wide variation in guideline adherence between different hospitals. Clearly, the consequences of substandard IUI care are considerable. Therefore, interventions that address the identified deficits in IUI care are warranted.

As time and resources for initiatives to change practice performance in health care are generally limited, targeted interventions to specifically improve aspects of care that are associated with better clinical outcomes are probably most effective and efficient (Grol and Grimshaw, 2003). However, up until now, data regarding the association between guideline adherence and outcome of infertility care were limited. This study revealed two statistically significant associations between guideline adherence and improved ongoing pregnancy rates after IUI which can direct future quality improvement activities. In other words, efforts should be made to specifically promote implementation of the guideline recommendations on sperm quality and the total number of IUI cycles.

Remarkably, adherence to the recommendation that HCG should be administered at a dose of 5000 IU in stimulated IUI was negatively associated with ongoing pregnancy after IUI. The lack of published evidence to support or oppose this association is surprising. The underlying reasons of this finding can be speculated upon. However, to judge this finding by its true merits, the causal relationship between the dose of HCG used in stimulated IUI and the ongoing pregnancy rate should be investigated in a randomized controlled trial first.

A statistically significant association between ≥90% guideline adherence and ongoing pregnancy after IUI could not be established for seven out of the 10 possible associations investigated. Similarly, the few studies previously performed found only small statistically significant or non-significant associations between guideline adherence and patient mortality rates for acute myocardial infarction, heart failure and pneumonia, despite large sample sizes (Fonarow et al., 2007; Horn, 2006; Werner and Bradlow, 2006). This raises a fundamental question: Why can't

adherence be successfully linked to guideline recommendations that are mostly based on solid research evidence to better patient outcomes in actual practice?

There are probably several explanations. An explanation might be that guideline recommendations are not always based on top-level evidence (Schouten et al., 2005). In the present study, 60% of the process indicators was supported by top-level evidence. Therefore, more well-designed studies, preferably randomized controlled trials, are needed to improve the strength of the supporting evidence of guideline recommendations. It is also possible that the findings about practice performance and clinical outcomes from homogeneous study samples that are used to form and support guideline recommendations may not be applicable to all patients in actual practice (Horn, 2006). Another, highly plausible explanation might be that the present study only investigated the association between guideline adherence and ongoing pregnancy and not all guideline recommendations are necessarily associated with improved ongoing pregnancy rates. For example, recommendations regarding the dose of gonadotrophins and cancellation criteria in stimulated IUI are associated with patient safety.

Although this study found mainly non-significant associations between guideline adherence and ongoing pregnancy after IUI, it cannot necessarily be said that guideline adherence is rather useless. Instead, attention should be focused on evaluating associations between guideline adherence and other outcomes of IUI care besides ongoing pregnancy, such as patient safety and cost effectiveness.

General considerations

Health-care delivery according to clinical practice guidelines is thought to be critical in achieving optimal outcomes. Reliable assessment of the extent to which practice performance in health care is consistent with guideline recommendations is therefore a key step in the dynamic process of quality assessment and improvement. This requires a set of valid guideline-based performance indicators. Tailored interventions to address shortcomings in practice performance come next. Preferably, the focus should be on care aspects strongly associated with improved outcomes. More research to establish such associations between guideline adherence and outcomes is necessary. To be sure that interventions to improve practice performance have the desired effect, ongoing assessment of practice performance in routine health care is warranted. To succeed, each quality control system will need adequate support and appropriate levels of funding. As all countries face common challenges to deliver appropriate, high-quality care to patients, there is an argument for more international collaboration in the development and research regarding clinical practice guidelines, performance indicators and implementation strategies to improve health care.

Acknowledgements

The authors thank the panel of 13 experts who contributed to the development of the indicators. They also thank Sabine van den Akker and Annelies Pellegrino for assisting with data collection and Jan van Doremalen and Reinier Akkermans for assisting with the statistical analyses. The contribution of the 10 Dutch hospitals for participating and making this study possible is gratefully acknowledged: Bernhoven Hospital—Oss, Bernhoven Hospital—Veghel, Canisius Wilhelmina Hospital—Nijmegen, Catharina Hospital—Eindhoven, Gelderse Vallei Hospital—Ede, Gelre Hospital—Apeldoorn, Radboud University Medical Centre—Nijmegen, Rijnstate Hospital—Arnhem, Slingeland Hospital—Doetinchem, St Antonius Hospital—Nieuwegein. This study was supported financially by the Netherlands Organisation for Health Research and Development, Grant No. 945-12-012, The Hague, The Netherlands.

References

Balen, A.H., Rutherford, A.J., 2007. Management of infertility. Br. Med. J. 335, 608—611.

Bensdorp, A.J., Cohlen, B.J., Heineman, M.J., et al., 2007. Intrauterine insemination for male subfertility. Cochrane Database Syst. Rev. (4). doi:10.1002/14651858.CD000360.pub4. Art. No. CD000360.

Boivin, J., Bunting, L., Collins, J.A., et al., 2007. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. Hum. Reprod. 22, 1506—1512.

Campbell, S.M., Braspenning, J., Hutchinson, A., et al., 2003. Research methods used in developing and applying quality indicators in primary care. Br. Med. J. 326, 816—819.

Campbell, S.M., Cantrill, J.A., Roberts, D., 2000. Prescribing indicators for UK general practice. Delphi consultation study. Br. Med. J. 321, 1—5.

Campbell, S.M., Roland, M.O., Middleton, E., et al., 2005. Improvements in quality of clinical care in English general practice 1998—2003: longitudinal observational study. Br. Med. J. 331, 1121.

Cantrill, J.A., Sibbald, B., Buetow, S., 1998. Indicators of the appropriateness of long-term prescribing in general practice in the United Kingdom: consensus development, face and content validity, feasibility, and reliability. Int. J. Qual. Health Care 7, 130—135.

Cousineau, T.M., Domar, A.D., 2007. Psychological impact of infertility. Best Pract. Res. Clin. Obstet. Gynaecol. 21, 293—308.

Danish Fertility Society (DFS), 2007. Clinical guideline No. 11: homologous intrauterine insemination. Available from: <http://www.fertilitetsselskab.dk/>.

Dickey, R.P., Taylor, S.N., Lu, P.Y., et al., 2005. Risk factors for high-order multiple pregnancy and multiple birth after controlled ovarian hyperstimulation: results of 4062 intrauterine insemination cycles. Fertil. Steril. 83, 671—683.

Dutch Society of Obstetrics and Gynaecology (NVOG), 2000. Clinical guideline No. 29: intrauterine insemination. Available from: <http://www.nvog.nl/files/iui.pdf/>.

The ESHRE Capri Workshop Group, 2009. Intrauterine insemination. Hum. Reprod. Update 15, 265—277.

Fauser, B.C., Devroey, P., Macklon, N.S., 2005. Multiple birth resulting from ovarian stimulation for subfertility treatment. Lancet 365, 1807—1816.

Fonarow, G.C., Abraham, W.T., Albert, N.M., et al., 2007. Association between performance measures and clinical outcomes for patients hospitalized with heart failure. JAMA 297, 61—70.

Goverde, A.J., McDonnell, J., Vermeiden, J.P., et al., 2000. Intrauterine insemination or in-vitro fertilisation in idiopathic subfertility and male subfertility: a randomised trial and cost-effectiveness analysis. Lancet 355, 13—18.

Greil, A.L., 1997. Infertility and psychological distress: a critical review of the literature. Soc. Sci. Med. 45, 1679—1704.

Grimes, D.A., Schulz, K.F., 2002. Bias and causal associations in observational research. Lancet 359, 248—252.

Grol, R., Baker, R., Moss, F., 2002. Quality improvement research: understanding the science of change in health care. Qual. Saf. Health Care 11, 110—111.

Grol, R., Grimshaw, J., 2003. From best evidence to best practice: effective implementation of change in patients' care. Lancet 362, 1225—1230.

Guzick, D.S., Carson, S.A., Coutifaris, C., et al., 1999. Efficacy of superovulation and intrauterine insemination in the treatment of infertility. N. Engl. J. Med. 340, 177—183.

Haagen, E.C., Hermens, R.P., Nelen, W.L., et al., 2006. Subfertility guidelines in Europe: the quantity and quality of intrauterine insemination guidelines. Hum. Reprod. 21, 2103—2109.

Horn, S.D., 2006. Performance measures and clinical outcomes. JAMA 296, 2731—2732.

Hughes, E.G., 2003. Stimulated intra-uterine insemination is not a natural choice for the treatment of unexplained subfertility. 'Effective treatment' or 'not a natural choice'? Hum. Reprod. 18, 912—914.

Luck, J., Peabody, J.W., Dresselhaus, T.R., et al., 2000. How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. Am. J. Med. 108, 642—649.

Lukassen, H.G., Schonbeck, Y., Adang, E.M., et al., 2004. Cost analysis of singleton versus twin pregnancies after in vitro fertilization. Fertil. Steril. 81, 1240—1246.

Mainz, J., 2003. Defining and classifying clinical indicators for quality improvement. Int. J. Qual. Health Care 15, 523—530.

McGlynn, E.A., Asch, S.M., Adams, J., et al., 2003. The quality of health care delivered to adults in the United States. N. Engl. J. Med. 348, 2635—2645.

Nachtigall, R.D., 2006. International disparities in access to infertility services. Fertil. Steril. 85, 871—875.

National Agency for Accreditation and Evaluation in Healthcare (ANAES), 1996. Clinical guideline No. 3: infertility of a couple. Le Concours Med. 40, 1—15.

National Institute for Health and Clinical Excellence (NICE), 2004. Clinical guideline: fertility: assessment and treatment for people with fertility problems, Chapter 10: intrauterine insemination. Available from: <http://www.nice.org.uk/nicemedia/pdf/CG011fullguideline.pdf/>.

Nelen, W.L., Hermens, R.P., Mourad, S.M., et al., 2007. Monitoring reproductive health in Europe: what are the best indicators of reproductive health? A need for evidence-based quality indicators of reproductive health care. Hum. Reprod. 22, 916—918.

Nyboe Andersen, A., Goossens, V., Bhattacharya, S., et al., 2009. Assisted reproductive technology and intrauterine inseminations in Europe, 2005: results generated from European registers by ESHRE. Hum. Reprod. 24, 1267—1287.

Rawal, N., Drakeley, A., Haddad, N., 2008. Intrauterine insemination practice in the UK. J. Obstet. Gynaecol. 28, 738—741.

Rubin, H.R., Pronovost, P., Diette, G.B., 2001. From a process of care to a measure: the development and testing of a quality indicator. Int. J. Qual. Health Care 13, 489—496.

Schouten, J.A., Hulscher, M.E., Wollersheim, H., et al., 2005. Quality of antibiotic use for lower respiratory tract infections at hospitals: (how) can we measure it? Clin. Infect. Dis. 41, 450—460.

Sheldon, T.A., Cullum, N., Dawson, D., et al., 2004. What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients' notes, and interviews. Br. Med. J. 329, 999.

Steures, P., van der Steeg, J.W., Hompes, P.G., et al., 2007. Intrauterine insemination in The Netherlands. Reprod. Biomed. Online 14, 110—116.

Steures, P., van der Steeg, J.W., Verhoeve, H.R., et al., 2004. Does ovarian hyperstimulation in intrauterine insemination for cervical factor subfertility improve pregnancy rates? Hum. Reprod. 19, 2263—2266.

Taylor, A., 2003. ABC of subfertility: extent of the problem. Br. Med. J. 327, 434—436.

Tummon, I.S., Asher, L.J., Martin, J.S., et al., 1997. Randomized controlled trial of superovulation and insemination for infertility associated with minimal or mild endometriosis. Fertil. Steril. 68, 8—12.

Verhaak, C.M., Smeenk, J.M., van Minnen, A., et al., 2005. A longitudinal, prospective study on emotional adjustment before, during and after consecutive fertility treatment cycles. Hum. Reprod. 20, 2253—2260.

Verhulst, S.M., Cohlen, B.J., Hughes, E., et al., 2006. Intra-uterine insemination for unexplained subfertility. Cochrane Database Syst. Rev. (4). doi:10.1002/14651858.CD001838.pub3. Art. No. CD001838.

Werner, R.M., Bradlow, E.T., 2006. Relationship between Medicare's hospital compare performance measures and mortality rates. JAMA 296, 2694—2702.

Declaration: The authors report no financial or commercial conflicts of interest.

Received 22 July 2009; refereed 13 August 2009; accepted 30 November 2009.