BMJ Open

Shared decision-making results in less variation between hospitals: a first, exploratory, study examining the relationship between shared decision-making and medical practice variation

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010894
Article Type:	Research
Date Submitted by the Author:	17-Dec-2015
Complete List of Authors:	Brabers, Anne; NIVEL Netherlands Institute for Health Services Research, van Dijk, Liset; NIVEL Netherlands institute for health services research, Groenewegen, Peter; NIVEL Netherlands institute for health services research; Utrecht University, Department of Sociology, Department of Human Geography van Peperstraten, Arno; Radboud University Medical Center, Department of Obstetrics and Gynaecology De Jong, Judith; NIVEL Netherlands institute for health services research,
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Health policy, Obstetrics and gynaecology
Keywords:	shared decision-making, medical practice variation, in vitro fertilisation, patients' preferences, patient involvement

Shared decision-making results in less variation between hospitals: a first, exploratory, study examining the relationship between shared decision-making and medical practice variation

Anne E.M. Brabers MSc[1§], Liset van Dijk PhD[1], Peter P. Groenewegen PhD[12], Arno M. van Peperstraten PhD[3], Judith D. de Jong PhD[1]

[1]NIVEL, the Netherlands Institute for Health Services Research, PO Box 1568, 3500 BN Utrecht, the Netherlands

[2]Department of Sociology, Department of Human Geography, Utrecht University, PO Box 80125, 3508 TC Utrecht, the Netherlands

[3]Department of Obstetrics and Gynaecology, Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, the Netherlands

[§]Corresponding author

Contact address corresponding author: PO Box 1568, 3500 BN Utrecht, the Netherlands, T: +31 (0)30 27 29 700, F: +31 (0)30 27 29 729

Email addresses:

AB: a.brabers@nivel.nl

LVD: l.vandijk@nivel.nl

PG: p.groenewegen@nivel.nl

AVP: arnovp@outlook.com

JDJ: j.dejong@nivel.nl

Word count: 4,153

Key-words shared decision-making, medical practice variation, in vitro fertilisation, patients' preferences, patient involvement

ABSTRACT

Objectives The hypothesis that shared decision-making (SDM) reduces medical practice variation is increasingly common, but no evidence is available. We aimed to elaborate further on this, and to perform a first, exploratory, analysis to examine this hypothesis. This analysis, based on a limited dataset, examined how SDM is associated with variation in the choice of single embryo transfer (SET) or double embryo transfer (DET) after in vitro fertilisation (IVF). We examined variation both between and within hospitals.

Design Secondary analysis RCT.

Setting Five hospitals in The Netherlands.

Participants 222 couples (woman aged <40) on a waiting list for a first IVF cycle, who could choose between SET and DET (i.e. ≥ 2 embryos available).

Interventions SDM via a multifaceted strategy aimed to empower couples in deciding how many embryos should be transferred. The strategy consisted of decision aid, support of IVF nurse, and the offer of reimbursement for an extra treatment cycle. Control group received standard IVF care.

Primary outcome measure Difference in variation due to SDM in the choice of SET or DET, both between and within hospitals.

Results There was large variation in the choice of SET or DET between hospitals in the control group. Lower variation between hospitals was observed in the group with SDM. Within most hospitals variation in the choice of SET or DET appeared to increase due to SDM. Particularly in hospitals where mainly DET was chosen in the control group.

Conclusions Although based on a limited dataset, our study gives a first insight that including patients' preferences through SDM results in less variation between hospitals, and indicates another pattern of variation within hospitals. Variation that results from patient preferences could be potentially named the informed patient rate. Our results provide the starting point for further research within this area.

Trial registration ClinicalTrials.gov NCT00315029.

ARTICLE SUMMARY

- This study is the first to elaborate further upon, and explore the association between, SDM and variation in medical treatment.
- Data from a RCT are used, which enables a comparison to be drawn between a situation with, and one without, SDM.
- A limitation is that we had only access to a limited dataset, and as such, we performed descriptive statistics to test our hypotheses.

INTRODUCTION

Considerable variation exists in medical treatment [1-4]. In a paternalistic model, the physician is the dominant actor deciding on this treatment [5 6]. This approach has been widely practiced, embedded in the idea that physicians decide on treatment based on both medical science and what is best for an individual patient [3], i.e. the belief "the doctor knows best". As such, physicians' professional judgements rather than patients' preferences often determined which treatment a patient received [3,p9 7]. As a result, variation was found to be related to physicians, rather than to patients [3 8 9]. In explaining variation, research therefore focused on the role of physicians, while patients' influence has received little attention [10]. In the past decades, however, the paternalistic model has been questioned. Also, the position of patients in health care has significantly altered. On an individual level they are supposed to take up an active role in their health [11], and are expected to be involved in decisions about their health [12]. There is, thus, an increased emphasis on including patients and their preferences in medical decision-making [13 14]. Since medical decision-making is decisive for variation in medical practice, it is questioned whether patients can still be ignored in theories about variation. Providing care that is respectful of, and responsive to, an individual patient's preferences, so-called patient-centredness, is regarded as of primary importance to health care alongside dimensions such as being safe, effective, timely, efficient, and equitable [15].

Medical decisions regarding treatment may change through the inclusion of patients' preferences as these preferences may deviate from physicians' professional judgements [16]. Including patients' preferences may result in different treatment choices, and patterns of variation. It has been hypothesized that patient involvement may reduce variation in medical practice, because research shows that patients, through a combination of education and participation, were less ready to accept certain procedures [17]. This also assumes that physicians are more diverse in their preferences than patients despite the fact that they have a shared training and socialisation that has no parallel among patients [17]. One specific approach to patient involvement is shared decision-making (SDM) [18 19]. SDM is defined as an approach where physicians and patients take decisions together using the best available evidence. Patients are helped to make informed choices by considering the options and the likely benefits and disadvantages of each option [20 21]. This is important as informed patients often prefer other treatments than their physician [3]. Research showed that, in general, informed patients prefer less invasive treatment options [3 22]. On the other hand, variation exists between physicians, since some of them prefer invasive treatments and others conservative treatments [23]. As such, it has been suggested that SDM – as a

case of patient involvement – reduces variation [19 24-26]. However, no clear evidence about the association between SDM and variation is available yet. There is no research which has identified exactly how or why SDM might reduce variation. Therefore, this study further elaborates upon the mechanisms that may explain why SDM reduces practice variation. In addition, we aim to perform a first, explorative, analysis on a limited dataset to examine the hypothesis that SDM reduces medical practice variation. Hereby, we make use of a clear example of a decision which depends on patients' preferences, the choice of either a single embryo transfer (SET) or double embryo transfer (DET) after in vitro fertilisation (IVF) [27].

SET prevents a multiple pregnancy with associated higher risks. DET results in higher live birth rates per treatment cycle [28 29] [see Box 1 for more information]. The percentage of SET varies considerably between countries [30-32]. For example, rates of SET ranged from 8.7 per cent in Moldova to 70.7 per cent in Sweden [32]. Likewise, major differences exist in how this complex decision is taken. These differences exist between countries, and between hospitals within the same country. In some hospitals the decision is based solely on clinical parameters, while in other hospitals patients are fully involved in the decision and physicians act as advisor [33]. If the physician decides for SET or DET, this decision is mainly based on physicians' professional judgements. As such, variation is likely. This can be explained by differences in opinions on, or enthusiasm for, certain procedures between individual physicians, and by differences in constraints and social influences for groups of physicians, for instance between hospitals [9 34-36]. When patients are involved, decisions may differ; informed patients often prefer less invasive treatments [3 22]. The RCT analysed in this study, in which SDM was used, concluded that educated couples, who understood the risks of twin pregnancies, were more inclined to choose SET. This is compared to couples receiving standard care [27].

The aim of this study is to examine how SDM is associated with variation in choosing SET or DET both **between** and **within** hospitals. We hypothesise that SDM is associated with less variation **between** hospitals, since we expect that, due to SDM, SET is chosen more often both in hospitals where physicians already preferred SET and in hospitals where physicians preferred DET, since educated couples prefer this. We also hypothesise that if DET is mainly preferred **within** a hospital, and there is thus hardly any variation, then SDM is expected to increase variation, because SET will be chosen more often due to SDM. Whereas, if SET is mainly preferred **within** a hospital, and there is thus hardly any variation, then SDM is expected not to change variation, since SET is still preferred due to SDM. Furthermore, we hypothesise that if DET and SET are both chosen **within** a hospital, and

there is thus large variation, SDM is expected to decrease variation, because, due to SDM, SET is likely to be chosen more often than DET.

METHODS

Description of the data

Data for this research were obtained from the RCT by Van Peperstraten et al. (2010) investigating the effect of a multifaceted empowerment strategy on the choice of a decision for SET or DET [27]. To empower couples to make this decision, Van Peperstraten et al. developed a multifaceted strategy comprising, an evidence-based decision aid (DA) [see Box 2 for more information], the support of an IVF nurse, and reimbursement for an additional cycle of IVF for couples for whom the choice of SET caused a reduced chance of pregnancy. The content of the DA and the reimbursement offer were discussed in person with a trained IVF nurse. All three elements of the strategy were provided before the counseling session that was part of standard care [27]. For further detailed information see Van Peperstraten et al. (2010). The control group received standard IVF care, including a session discussing the choice of SET or DET. Next to this standard care, the intervention group received the multifaceted empowerment strategy [27]. Before the study, in 2005, 39% of the couples underwent SET after the first cycle [37]. The RCT was performed in five hospitals in the Netherlands. It included couples on the IVF waiting list between November 2006 and July 2007. The follow-up was continued until December 2008. Couples of women under 40 were included if they were on the waiting list for their first IVF cycle ever or a first cycle after a previous successful IVF. Couples were excluded if SET was mandatory due to a strict medical indication. Written informed consent was provided by the couples before participation [27].

Selection of the data

In total, 308 couples at the beginning of their first IVF cycle were included in the intention to treat analysis (ITT) of Van Peperstraten et al. (2010) [27]. In all five hospitals approximately half of the couples received standard care, while the other half received the intervention. In this study, only couples that had the opportunity to choose between SET and DET were included. We, therefore, omitted from the 308 couples included in the ITT all couples: 1) where the woman was pregnant before starting IVF (N=20); 2) that never started IVF (N=13); and 3) that had none or just one embryo available (N=39 respectively N=14). Our sample included 222 couples, 113 in the control group and 109 in the intervention group, respectively. The outcome measure used in this study was the choice of either SET or DET. The data on this outcome were collected by Van Peperstraten et al. (2010) from

local IVF registries [27]. Other variables included were whether a couple was involved in the intervention group or in the control group, and the hospital in which a couple was treated. In addition, we included four variables that are of medical relevance and might, therefore, affect the choice of SET or DET, and thus practice variation. For example, the older the woman is, the less likely she will become pregnant and the more likely she will have twins. The four variables included were: 1) the age of the woman (in years); 2) the duration of infertility (in years); 3) the presence of a good quality embryo (yes/no); and 4) any previous pregnancies (yes/no). Data on the presence of a good quality embryo were collected by Van Peperstraten et al. (2010) from local IVF registries [27]. Data for the other three variables were collected through a patient questionnaire which couples received when included to the study [27]. The woman's age and duration of infertility were calculated in this study on 31 December 2008 (= end of follow-up).

Statistical analyses of the data

We examined whether the control and intervention group were comparable for the characteristics included by performing descriptive statistics, and chi^2 tests (categorical variables) and t-tests (continuous variables) ($p < 0.05$). We then examined whether the five hospitals included did significantly differ with respect to the four variables that are of medical relevance. If there were differences between the five hospitals then we had to take these into account throughout the rest of our analyses, since these may have an impact on the choice of SET or DET. We performed descriptive statistics per hospital for the four variables. By chi^2 tests (categorical variables) and one-way analyses of variance (ANOVA) (continuous variables), we tested if there were significant differences between the five hospitals for the woman's age, the duration of infertility, the presence of a good quality embryo, and for previous pregnancies ($p < 0.05$). If a significant difference was found between the hospitals for one of the aforementioned variables, we then performed an additional analysis to examine if there was an association between that variable and the outcome measure.

We then calculated, for each hospital, the percentage of couples that chose SET or DET, both in the control and in the intervention groups. We examined this in order to confirm that educated couples are inclined to choose SET. Next, we examined the variation **between** the hospitals. We first calculated for each hospital the percentage of SET in the control group and then in the intervention group. Thereafter, we calculated the range of SET percentages for the control groups and the intervention groups. A smaller range or difference between the highest and the lowest percentage of SET, implies less variation. Thus, if SDM is associated with less variation between hospitals, then the range of SET percentages for the intervention group is smaller than that for the control group.

We now examined the variation **within** the hospitals by looking at the differences in variation between the control and the intervention group in each hospital. We calculated for each hospital the absolute difference between SET and DET in the control as well as in the intervention group. For example, if 40% chose SET and 60% DET, the absolute difference is 20. There is no variation if the proportion of SET to DET is 0% compared to 100%, or vice versa. Thus, there is no variation if the absolute difference is 100. On the other hand, the most variation is observed if the proportion of SET to DET is 50% to 50%. Thus, there is an absolute difference of 0. We can thus create a scale ranging from 0 to 100, where a score closer to 100 means less variation. We compared the scores of the control and the intervention group for each hospital. If the score in a hospital is higher in the intervention group than in the control group, and thus closer to 100 (= no variation), then SDM is associated with less variation within that hospital. Complementary to the descriptive statistics, we performed a multilevel analysis (MLA) in MLwiN to examine variation between hospitals. A MLA takes into account the nested structure of the data as well as the differences in the number of patients per hospital. All statistical analyses were carried out using STATA, version 13.1.

RESULTS

Characteristics of the couples included

A description of the 222 couples included is given in Table 1. The number of couples included ranged from twelve couples in hospital five to 153 couples in hospital one (see Table 1). The control and intervention group did not differ significantly with respect to the characteristics included. Furthermore, no significant differences were observed between the five hospitals for the mean duration of infertility (p=0.256), the presence of a good quality embryo (p=0.406), and for previous pregnancies (p=0.403) (see Table 2). ANOVA showed a significant difference (p=0.032) for the variable, woman's age. An additional t-test showed no difference between woman's age and the choice of SET or DET (p=0.346). Thus, we decided not to include these four variables throughout the rest of the analyses.

Choice of SET

Table 3 shows the numbers and percentages of SET and DET for both the control and intervention groups, in total and per hospital. In total, 52% of the couples included in the intervention group chose SET, in comparison with 39% of the couples in the control group (p=0.046). To be more specific, in four of the five hospitals, couples in the intervention groups more often chose SET than DET. In hospital four, however, couples in the intervention

group more often chose DET than SET (80% vs. 20%). Although couples in the intervention group more often chose SET (20%) than couples in the control group (0%).

Variation between hospitals

The range of SET in both the control and intervention groups can also be observed in Table 3. The percentages of SET in the control groups ranged from 0.0% to 85.7%, while the percentages of SET in the intervention groups ranged from 20.0% to 87.5%. Therefore the range of SET is smaller in the intervention group than in the control group, which is an indication that SDM reduced variation in the choice of SET or DET between hospitals. The MLA also indicated that the variation between hospitals was lower in the intervention group than in the control group. However, the difference was not significant.

Variation within hospitals

Figure 1 shows the differences in variation within hospitals by illustrating, per hospital, the absolute difference between SET and DET in the control group (no SDM) and the intervention group (SDM). In one hospital (number 2) the absolute difference in the control group and the intervention group is the same. This means that the variation within hospital 2 is the same with or without SDM. Within hospital 3, SDM appears to be associated with less variation, since the absolute difference in the control group is lower than in the intervention group (14 and 75, respectively). On the other hand, within the other three hospitals, numbers 1, 4 and 5, SDM appears to be associated with more variation. Within these three hospitals the absolute difference in the control group is higher than in the intervention group (see Figure 1). Therefore, within some of the hospitals included, SDM appears to be associated with more variation, while within other hospitals SDM appears to be associated with less or the same level of variation.

DISCUSSION AND CONCLUSIONS

Principal findings

This study further elaborated upon and explored the association between SDM and variation in the choice of SET or DET both between and within hospitals. There was large variation in the choice of SET or DET between hospitals in the control group. Lower variation between hospitals was observed in the group with SDM. Furthermore, we observed that within most hospitals the variation in the choice of SET or DET appeared to increase due to SDM. This was particularly in hospitals where mainly DET was chosen in the control group.

What this study adds

Literature suggests that SDM reduces variation [19 24-26]. There was however, up to now, no clear evidence about this association. This study is the first that explored this association based on a case concerning the choice of SET or DET after IVF. We noticed that SDM reduces variation between hospitals, while the variation within most hospitals appears to increase. The hypothesis in literature that SDM reduces variation is based on the observation that informed patients more often prefer less invasive treatments [3 22]. We found that in most hospitals couples in the intervention group more often chose SET. Although this does not imply that there will be less variation, since our results indicate that variation within most hospitals increased. This is because the level of variation without SDM differed between hospitals. For example, in some of the hospitals included mainly DET was preferred and there was thus almost no variation. Due to SDM, however, SET was chosen more often, and thus the variation increased within such hospitals, since now both SET and DET are chosen. A subsequent implication is that an overall decrease in variation between hospitals, provides no indication about the change in variation within an individual hospital. Although based on a limited dataset, this study gives a first insight that SDM results in less variation between hospitals while suggesting another pattern of variation within hospitals.

Further research

This research focused on just one case study, and had only access to a limited dataset. The results, therefore, have to be interpreted with caution and further research is necessary both to underpin our results and to examine questions that remain unanswered. Nevertheless, our study provides a starting point for further empirical research within this area. For many medical problems no absolute best treatment option is available and so there are significant trade-offs among the available options [38 39]. We expect, however, that our results apply generally to medical problems with no absolute best treatment option. Decisions concerning such problems are defined as preference sensitive, since they depend on considerations of the benefits, disadvantages, and uncertainties of each treatment. For example some patients will prefer to accept a small risk of death in order to attempt to improve their function, while others will not [19 39]. Therefore, the best decisions cannot be made without including patients' preferences [38 39]. Well-known examples include chronic back pain, early-stage breast cancer and prostate cancer. For examples such as these it is believed that variation will change as a result of SDM. Future research has to confirm this by making use of data from multicentre RCT studies that applied SDM and/or decision aids as an intervention. Such RCT studies have been carried out [22], but have focused, comparable to

the study we used, on outcome measures other than variation. Therefore, we decided in this study to perform a secondary analysis. Any possible multicentre studies should include a control and an intervention group which could thus measure actual treatment choices with and without SDM. This would allow researchers to examine whether SDM changes the pattern of variation, by, for example, using the same method as we did.

Our results show that SDM results in less variation between hospitals and indicates another pattern of variation within hospitals, confirming our hypotheses. These results appear to show that the decisions made by informed patients have a pattern too. Choices made by informed patients appear to have a rate which deviates from baseline rates, irrespective of whether those are "low" or "high". This could be potentially named the informed patient rate. However, it can be questioned whether the rates we observed are indeed the informed patient rate, that is the results of what the couples want. It is possible that not all patients were able, or preferred, to take a shared decision about the choice of SET or DET. Moreover, it can be questioned to what extent, in SDM interventions and in decision aids in general, patients' preferences can come into their own, or whether such an intervention is a reflection of the preferences of the physicians who developed it. Further research has to examine whether the actual choice was indeed the patients' preference and whether there are differences between groups of patients in this. In addition, a different pattern of variation due to SDM might be a positive indication for the quality of care. Good health care requires, among others, providing care that is respectful of, and responsive to, individual patients' preferences [15]. This is particularly true for preference sensitive decisions, since these decisions depend on patients' preferences regarding the benefits, disadvantages and uncertainties of each treatment. Further research has to examine whether SDM results in better quality of care for preference sensitive decisions.

The broad context of this study is about the influence of patients and their preferences on variation in medical practice. SDM is one option for including patients' preferences in medical decision-making. There are other options through which patients can express their preferences, and thus to influence the pattern of variation. For example, patients differ in how much pressure they are able to put on physicians [35 40]. They differ in their ability to take part in discussions over treatment with their physicians. Some patients are expected to be able to ask their physician for another treatment than, for example, the treatment that is recommended in a guideline or the one that is preferred by the physician [35 41]. If this is the case, then patients' preferences appear to influence the treatment chosen, and thus the variation. Further research is recommended into these situations.

The strengths and weaknesses of the study

A strength of our study is that we are the first to elaborate further upon and explore the association between SDM and variation in medical treatment. We examined this association to get insight into whether including patients' preferences through SDM results in another pattern of variation in medical treatment. Another strength is the use of data from a RCT. We had the opportunity to compare the variation in the choice of SET or DET with, and without, SDM. It might be possible that the choice of SET or DET in the control group is influenced by physicians, since they treated both couples in the intervention and in the control group. Ideally, data would have been available about the percentage of SET and DET before the RCT, allowing us to compare the intervention and control group with these percentages. Though another study showed that in 2005, before the RCT, 39% of twin prone couples in two Dutch hospitals chose SET [37], which is comparable to the percentage of SET in the control group. It seems plausible to use the control group as the situation before SDM. We performed descriptive statistics to analyse our data, because of the low numbers of couples included in the hospitals. We have taken into account the nested structure of the data by performing our analyses per hospital. However, we did not take into account the differences in the number of patients per hospital. We therefore also performed an MLA to examine the variation between hospitals. The MLA supported the results of the descriptive statistics, however, the difference was not significant. From the dataset it was also known only in which hospital a couple was treated, but not by which physician within that hospital. However, only one or two physicians per hospital treated all couples in that hospital so we do not expect that this will affect our conclusions. Further research should have access to a larger dataset, preferably using multilevel analysis in order to test the hypothesis of this study. This would acknowledge that patients are nested hierarchically within physicians and physicians within hospitals. A final limitation might be that the intervention consisted of different elements, and thus it is difficult to assess separately the effects of these elements. Despite the fact that at the end of the follow-up period only 4% of the couples qualified for reimbursement of a fourth cycle, reimbursement might have played a role in the decision [27].

Conclusions

This study was the first to elaborate further upon and explore the relationship between including patients' preferences in medical decision-making and practice variation. Although based on a limited dataset, our study gives a first insight that including patients' preferences through SDM results in less variation between hospitals, and indicates another pattern of variation within hospitals. The variation that results from patient preferences

could be potentially named the informed patient rate. The results of this study provide the starting point for further empirical research within this area.

Acknowledgements We thank Peter Spreeuwenberg (NIVEL) for advising on and performing the MLA.

Authorship AB was involved in the design of the study, performed the statistical analyses, and drafted the manuscript. JDJ and LVD were involved in the design of the study, assisted in interpreting the results, assisted in drafting the manuscript, and revised the manuscript. PG was involved in the design of the study, and revised the manuscript. AVP was the primary author of the original RCT and was therefore responsible for the data collection, and revised the manuscript. All authors have read and approved the final manuscript.

Conflict of interest All authors have completed the unified competing interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare (1) all no financial support for the work submitted from anyone other than their employer; (2) AB, PG, AVP and JDJ no financial relationships with commercial entities that might have an interest in the work submitted. LVD received unrestricted grants from Bristol-Myers Squibb, Pfizer and Astra Zeneca for studies not related to this study; (3) all no spouses, partners, or children with relationships with commercial entities that might have an interest in the work submitted; and (4) all no non-financial interests that may be relevant to the submitted work.

Source of funding The RCT study was funded by the Netherlands Organisation for Health Research and Development (grant no 945-16-105). For this secondary analysis of the RCT no additional funding was obtained.

Ethics approval The RCT study was approved by the regional ethics committee for medical research.

Data sharing statement The dataset of the RCT study is available upon request from AVP.

Transparency All authors are a guarantor and affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no imported aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

REFERENCES

1. Paul-Shaheen P, Clark JD, Williams D. Small area analysis: a review and analysis of the North American literature. Journal of Health Politics, Policy and Law 1987;**12**(4):741-809.

2. Wennberg JE. Understanding geographic variations in health care delivery. NEJM 1999;**340**(1):52-3.

3. Wennberg JE. *Tracking medicine: a researcher's quest to understand health care*: Oxford University Press, 2010.

4. Corallo AN, Croxford R, Goodman DC, et al. A systematic review of medical practice variation in OECD countries. Health Policy 2014;**114**(1):5-14.

5. Emanuel EJ, Emanuel LL. Four models of the physician-patient relationship. JAMA 1992;**267**(16):2221-6.

6. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med 1997;**44**(5):681-92.

7. Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. BMJ 2002;**325**(7370):961.

8. McPherson K. Why do variations occur? In: Andersen TF, Mooney G, eds. The challenges of medical practice variations. London: The Macmillan Press, 1990:16-35.

9. Wennberg JE, Gittelsohn A. Health care delivery in Maine I: patterns of use of common surgical procedures. J Maine Med Assoc 1975;**66**(5):123-30, 49.

10. Mercuri M, Gafni A. Medical practice variations: what the literature tells us (or does not) about what are warranted and unwarranted variations. J Eval Clin Pract 2011;**17**(4):671-7.

11. van de Bovenkamp HM, Trappenburg MJ, Grit KJ. Patient participation in collective healthcare decision making: the Dutch model. Health Expect 2010;**13**(1):73-85.

12. Guadagnoli E, Ward P. Patient participation in decision-making. Soc Sci Med 1998;**47**(3):329-39.

13. Frosch DL, Kaplan RM. Shared decision making in clinical medicine: past research and future directions. Am J Prev Med 1999;**17**(4):285-94.

14. Fraenkel L. Incorporating patients' preferences into medical decision making. Med Care Res Rev 2013;**70**(1 Suppl):80s-93s.

15. Institute of Medicine. *Crossing the quality chasm: A new health system for the 21st century*: National Academies Press, 2001.

16. Wennberg JE. *Variation in Use of Medicare Services among Regions and Selected Academic Medical Centers: Is More Better?*: Commonwealth Fund, 2005.

17. Greer AL, Goodwin JS, Freeman JL, et al. Bringing the patient back in. Guidelines, practice variations, and the social context of medical practice. Int J Technol Assess Health Care 2002;**18**(4):747-61.

18. Charles C, Gafni A, Whelan T. Decision-making in the physician-patient encounter: revisiting the shared treatment decision-making model. Soc Sci Med 1999;**49**(5):651-61.

19. O'Connor AM, Llewellyn-Thomas HA, Flood AB. Modifying unwarranted variations in health care: shared decision making using patient decision aids. Health Affairs 2004:Var63-72.

20. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012;**27**(10):1361-7.

21. Elwyn G, Laitner S, Coulter A, et al. Implementing shared decision making in the NHS. BMJ 2010;**341**:c5146.

22. Stacey D, Bennett CL, Barry MJ, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2011(10):Cd001431.

23. McClure W. Toward development and application of a qualitative theory of hospital utilization. Inquiry 1982;**19**(2):117-35.

24. O'Connor AM, Wennberg JE, Legare F, et al. Toward the 'tipping point': decision aids and informed patient choice. Health Affairs 2007;**26**(3):716-25.

25. Stiggelbout AM, Van der Weijden T, De Wit MP, et al. Shared decision making: really putting patients at the centre of healthcare. BMJ 2012;**344**:e256.

26. Weinstein JN, Bronner KK, Morgan TS, et al. Trends and geographic variations in major surgery for degenerative diseases of the hip, knee, and spine. Health Affairs 2004:Var81-9.

27. van Peperstraten A, Nelen W, Grol R, et al. The effect of a multifaceted empowerment strategy on decision making about the number of embryos transferred in in vitro fertilisation: randomised controlled trial. BMJ 2010;**341**:c2501.

28. Pandian Z, Marjoribanks J, Ozturk O, et al. Number of embryos for transfer following in vitro fertilisation or intra-cytoplasmic sperm injection. Cochrane Database Syst Rev 2013;**7**:Cd003416.

29. Pinborg A. IVF/ICSI twin pregnancies: risks and prevention. Hum Reprod Update 2005;**11**(6):575-93.

30. Andersen AN, Goossens V, Ferraretti AP, et al. Assisted reproductive technology in Europe, 2004: results generated from European registers by ESHRE. Hum Reprod 2008;**23**(4):756-71.

31. de Mouzon J, Goossens V, Bhattacharya S, et al. Assisted reproductive technology in Europe, 2006: results generated from European registers by ESHRE. Hum Reprod 2010;**25**(8):1851-62.

32. Ferraretti AP, Goossens V, Kupka M, et al. Assisted reproductive technology in Europe, 2009: results generated from European registers by ESHRE. Hum Reprod 2013;**28**(9):2318-31.

33. van Peperstraten A. Implementation of single embryo transfer: a patient directed strategy. Radboud Universiteit Nijmegen, 2009.

34. Chassin MR. Explaining geographic variations. The enthusiasm hypothesis. Med Care 1993;**31**(5 Suppl):Ys37-44.

35. De Jong JD. Explaining medical practice variation. Social organization and institutional mechanisms. NIVEL, 2008.

36. Westert GP, Groenewegen PP. Medical practice variations: changing the theoretical approach. Scand J Public Health 1999;**27**(3):173-80.

37. Van Peperstraten AM, Kreuwel IA, Hermens RP, et al. Determinants of the choice for single or double embryo transfer in twin prone couples. Acta Obstet Gynecol Scand 2008;**87**(2):226-31.

38. Barry MJ. Shared decision making: informing and involving patients to do the right thing in health care. J Ambul Care Manage 2012;**35**(2):90-8.

39. Dartmouth Atlas. A Dartmouth Atlas Project Topic Brief Preference-Sensitive Care. Lebanon: The Dartmouth Atlas of Health Care, 2007.

40. Westert GP, Groenewegen PP, Lindenberg SM. De invloed van ziekenhuispatiënten op de duur van de hospitalisatie. TSG 1991;**69**:184-94.

41. Hajjaj FM, Salek MS, Basra MK, et al. Non-clinical influences on clinical decision-making: a major challenge to evidence-based practice. J R Soc Med 2010;**103**(5):178-87.

42. No authors listed. Prevention of twin pregnancies after IVF/ICSI by single embryo transfer. ESHRE Campus Course Report. Hum Reprod 2001;**16**(4):790-800.

43. Wolner-Hanssen P, Rydhstroem H. Cost-effectiveness analysis of in-vitro fertilization: estimated costs per successful pregnancy after transfer of one or two embryos. Hum Reprod 1998;**13**(1):88-94.

44. Lukassen HG, Schonbeck Y, Adang EM, et al. Cost analysis of singleton versus twin pregnancies after in vitro fertilization. Fertil Steril 2004;**81**(5):1240-6.

45. van Peperstraten AM, Hermens RP, Nelen WL, et al. Deciding how many embryos to transfer after in vitro fertilisation: development and pilot test of a decision aid. Patient Educ Couns 2010;**78**(1):124-9.

46. Elwyn G, O'Connor A, Stacey D, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ 2006;**333**(7565):417.

BOXES

Box 1: The choice of single or double embryo transfer: A complex decision-making problem

The choice of a single or double embryo transfer after IVF is a complex decision-making problem because of the need to find a balance between the risk of complications of multiple birth and the best chance of pregnancy [27]. Some sub-fertile couples and professionals regard twin pregnancies as a success, however, they could also be considered as a side effect or even a complication [42]. Twin pregnancies are associated with higher morbidity and mortality rates for both mother and child compared to singleton pregnancies [29]. Moreover, complications of twin pregnancies cause substantial use of medical budgets [43 44]. Subsequently, twin pregnancies are increasingly regarded as undesirable. To prevent twin pregnancies professionals and couples could choose single embryo transfer (SET) instead of double embryo transfer (DET) [29 42]. However, this may be disadvantageous, since it could result in a lower pregnancy rate per IVF cycle [28]. The choice of SET or DET is ideally decided through SDM [45].

Box 2: The choice of single or double embryo transfer: An evidence-based decision aid

Van Peperstraten et al. (2010) developed and tested the evidence-based decision aid (DA) for deciding how many embryos to transfer during IVF [45]. The DA was developed according to the checklist of the International Patient Decision Aids Standards Collaboration, which consists of 50 items divided between three domains, content, development, and effectiveness [45 46]. The purpose of the DA is to give couples all the information needed to make the choice to transfer one or two embryos and to relate the information to their own personal situation. The DA consists of three chapters: 1) information about the chances of a single pregnancy or a twin pregnancy; 2) information about the risks of twin pregnancies; and 3) an explanation of the available options and an action plan [45]. The DA is available in English at: www.umcn.nl/ivfda-en.

TABLES AND FIGURES

Table 1. Characteristics of the couples included (N = 222). Values are numbers unless otherwise stated.

Characteristics	Control group (N = 113)	Intervention group (N = 109)	Total (N = 222)	p-value[c]
Hospital				
hospital 1	79	74	153	N.A.
hospital 2	7	7	14	
hospital 3	7	8	15	
hospital 4	13	15	28	
hospital 5	7	5	12	
Mean (sd) age of woman (years)[a]	33.9 (3.85) (range 21-41 years)	33.5 (3.88) (range 25-41 years)	33.7 (3.86) (range 21-41 years)	0.475
Mean (sd) duration of infertility (years)[a]	4.03 (2.08) (range 1-13 years) (N=101)	3.94 (1.91) (range 1-12 years) (N=98)	3.98 (2.00) (range 1-13 years) (N=199)	0.749
Presence of a good quality embryo				
no	41	28	69	0.088
yes	72	81	153	
Previous pregnancies[b]	(N=113)	(N=108)	(N=221)	
no	63	63	126	0.698
yes	50	45	95	

[a] Calculated on December 31, 2008 based on information filled out in patients' questionnaires. As a result, we have a higher mean for age and duration of infertility than Van Peperstraten et al., BMJ, 2010.

[b] Based on the question: "Have you ever been pregnant?"

[c] $p < 0.05$ is significant

Table 2. Characteristics of the couples included per hospital.

Characteristics	Total (N = 222)	p-value[c]
Mean (sd) age of woman (years)[a]	33.7 (3.86) (21-41)	
hospital 1 (N = 153)	33.8 (3.63) (21-41)	**0.032**
hospital 2 (N = 14)	30.9 (4.54) (25-39)	
hospital 3 (N = 15)	34.6 (3.96) (28-40)	
hospital 4 (N = 28)	33.4 (4.53) (25-41)	
hospital 5 (N = 12)	35.2 (2.86) (30-38)	
Mean (sd) duration of infertility (years)[a]	4.0 (2.00) (1-13)	
hospital 1 (N = 139)	4.1 (2.20) (1-13)	0.256
hospital 2 (N = 11)	4.0 (1.26) (2-6)	
hospital 3 (N = 11)	4.3 (1.85) (2-8)	
hospital 4 (N = 26)	3.2 (1.22) (1-6)	
hospital 5 (N = 12)	4.0 (1.13) (2-6)	
Presence of a good quality embryo (% yes)	68.9%	
hospital 1 (N = 153)	72.6%	0.406
hospital 2 (N = 14)	64.3%	
hospital 3 (N = 15)	66.7%	
hospital 4 (N = 28)	60.7%	
hospital 5 (N = 12)	50.0%	
Previous pregnancies[b] (% yes)	43.0%	
hospital 1 (N = 153)	45.8%	0.403
hospital 2 (N = 14)	21.4%	
hospital 3 (N = 14)	50.0%	
hospital 4 (N = 28)	39.3%	
hospital 5 (N = 12)	33.3%	

[a] Calculated on December 31, 2008 based on information filled out in patients' questionnaires. As a result, we have a higher mean for age and duration of infertility than Van Peperstraten et al., BMJ, 2010.

[b] Based on the question: "Have you ever been pregnant?"

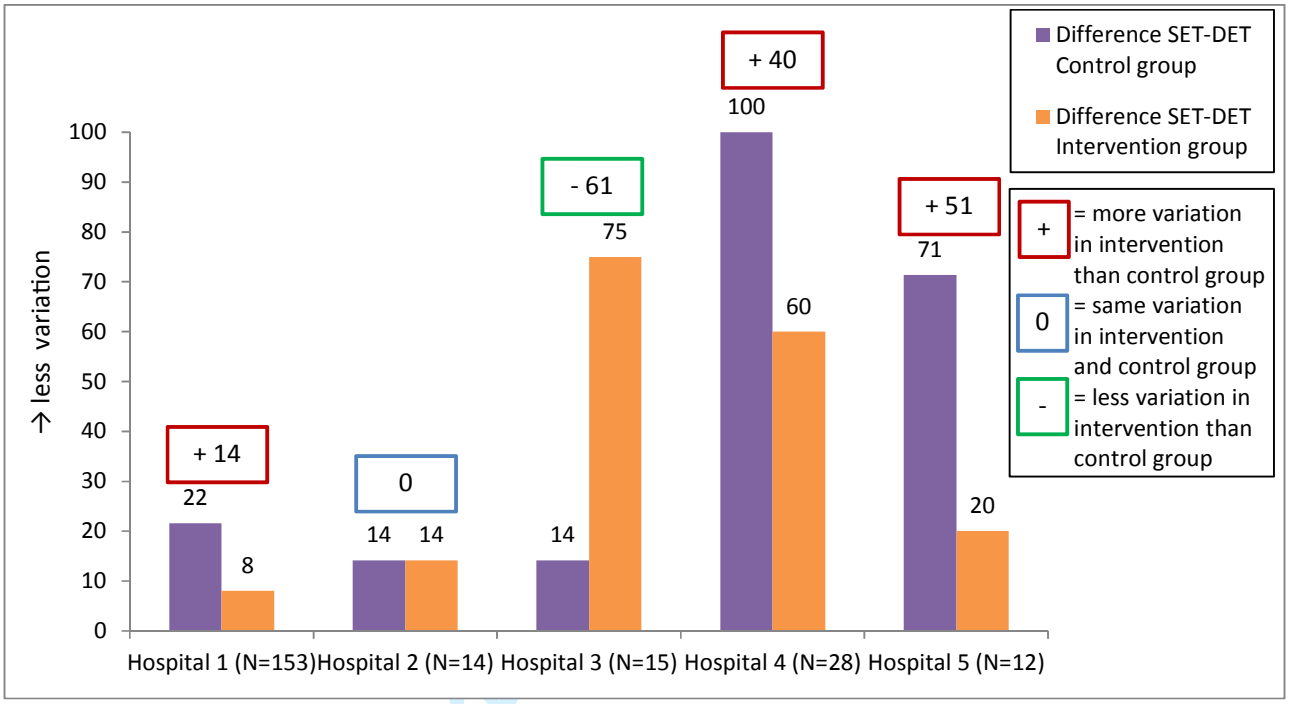
[c] $p < 0.05$ is significant

Table 3. The choice of SET or DET total group, and per hospital. Values are numbers (%) unless otherwise stated.

	Control group	Intervention group	p-value[a]
Total			
SET	44 (38.9%)	57 (52.3%)	**0.046**
DET	69 (61.1%)	52 (47.7%)	
Hospital 1			
SET	31 (39.2%)	40 (54.1%)	0.066
DET	48 (60.8%)	34 (46.0%)	
Hospital 2			
SET	3 (42.9%)	4 (57.1%)	0.593
DET	4 (57.1%)	3 (42.9%)	
Hospital 3			
SET	4 (57.1%)	7 (87.5%)	0.185
DET	3 (42.9%)	1 (12.5%)	
Hospital 4			
SET	0 (0.0%)	3 (20.0%)	0.088
DET	13 (100.0%)	12 (80.0%)	
Hospital 5			
SET	6 (85.7%)	3 (60.0%)	0.310
DET	1 (14.3%)	2 (40.0%)	

[a] $p < 0.05$ is significant

Figure 1: Variation within hospitals. A measure of variation for the control and intervention groups per hospital.



CONSORT 2010 checklist of information to include when reporting a randomised trial*

We performed a secondary analysis of a RCT. The original RCT was published in the BMJ (2010) as "The effect of a multifaceted empowerment strategy on decision making about the number of embryos transferred in in vitro fertilisation: randomised controlled trial" by Van Peperstraten et al. We included this article as supplement to our submission.

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	See Van Peperstraten et al. (2010)
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	Page 2 of this submission
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	Page 4-5 of this submission
	2b	Specific objectives or hypotheses	Page 4-5 of this submission
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	See Van Peperstraten et al. (2010)
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	NA
Participants	4a	Eligibility criteria for participants	See Van Peperstraten et al. (2010) and Page 6 of this submission
	4b	Settings and locations where the data were collected	See Van Peperstraten et al. (2010) and Page 6 of this submission
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	See Van Peperstraten et al. (2010) and Page 6 of this submission
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	Page 6-7 of this submission

	6b	Any changes to trial outcomes after the trial commenced, with reasons	NA
Sample size	7a	How sample size was determined	See Van Peperstraten et al. (2010) en Page 6 of this submission
	7b	When applicable, explanation of any interim analyses and stopping guidelines	NA
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	See Van Peperstraten et al. (2010)
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	See Van Peperstraten et al. (2010)
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	See Van Peperstraten et al. (2010)
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	See Van Peperstraten et al. (2010)
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	NA
	11b	If relevant, description of the similarity of interventions	NA
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	Page 7-8 of this submission
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	NA
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	Table 1 and 2 and 3 of this submission
	13b	For each group, losses and exclusions after randomisation, together with reasons	Page 6 of this submission
Recruitment	14a	Dates defining the periods of recruitment and follow-up	See Van Peperstraten et al. (2010) and Page 6 of this submission
	14b	Why the trial ended or was stopped	NA
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Table 1 and 2 of this submission
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Table 1 and 2 and 3 of this submission
Outcomes and	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its	Table 3 and Figure 1 and

estimation		precision (such as 95% confidence interval)	Page 8-9 of this submission
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	Table 3 and Figure 1
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	NA
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	NA
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	See Van Peperstraten et al. (2010) and Page 12 of this submission
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Page 10 of this submission
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	Page 10-12 of this submission
Other information			
Registration	23	Registration number and name of trial registry	See Van Peperstraten et al. (2010) and Page 2 of this submission
Protocol	24	Where the full trial protocol can be accessed, if available	NA
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	See Van Peperstraten et al. (2010) and Page 13 of this submission

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

BMJ Open

Does a strategy to promote SDM reduce medical practice variation in the choice of either single or double embryo transfer after IVF? A secondary analysis of an RCT.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010894.R1
Article Type:	Research
Date Submitted by the Author:	25-Feb-2016
Complete List of Authors:	Brabers, Anne; NIVEL Netherlands Institute for Health Services Research, van Dijk, Liset; NIVEL Netherlands institute for health services research, Groenewegen, Peter; NIVEL Netherlands institute for health services research; Utrecht University, Department of Sociology, Department of Human Geography van Peperstraten, Arno; Radboud University Medical Center, Department of Obstetrics and Gynaecology De Jong, Judith; NIVEL Netherlands institute for health services research,
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Health policy, Obstetrics and gynaecology
Keywords:	shared decision-making, medical practice variation, in vitro fertilisation, patients' preferences, patient involvement

Does a strategy to promote SDM reduce medical practice variation in the choice of either single or double embryo transfer after IVF? A secondary analysis of an RCT.

Anne E.M. Brabers MSc[1§], Liset van Dijk PhD[1], Peter P. Groenewegen PhD[12], Arno M. van Peperstraten PhD[3], Judith D. de Jong PhD[1]

[1]NIVEL, the Netherlands Institute for Health Services Research, PO Box 1568, 3500 BN Utrecht, the Netherlands

[2]Department of Sociology, Department of Human Geography, Utrecht University, PO Box 80125, 3508 TC Utrecht, the Netherlands

[3]Department of Obstetrics and Gynaecology, Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, the Netherlands

[§]Corresponding author

Contact address corresponding author: PO Box 1568, 3500 BN Utrecht, the Netherlands, T: +31 (0)30 27 29 700, F: +31 (0)30 27 29 729

Email addresses:

AB: a.brabers@nivel.nl

LVD: l.vandijk@nivel.nl

PG: p.groenewegen@nivel.nl

AVP: arnovp@outlook.com

JDJ: j.dejong@nivel.nl

Word count: 4,547

Key-words shared decision-making, medical practice variation, in vitro fertilisation, patients' preferences, patient involvement

ABSTRACT

Objectives The hypothesis that shared decision-making (SDM) reduces medical practice variation is increasingly common, but no evidence is available. We aimed to elaborate further on this, and to perform a first, exploratory, analysis to examine this hypothesis. This analysis, based on a limited dataset, examined how SDM is associated with variation in the choice of single embryo transfer (SET) or double embryo transfer (DET) after in vitro fertilisation (IVF). We examined variation between and within hospitals.

Design A secondary analysis of an RCT.

Setting Five hospitals in The Netherlands.

Participants 222 couples (woman aged <40) on a waiting list for a first IVF cycle, who could choose between SET and DET (i.e. ≥ 2 embryos available).

Intervention SDM via a multifaceted strategy aimed to empower couples in deciding how many embryos should be transferred. The strategy consisted of decision aid, support of IVF nurse, and the offer of reimbursement for an extra treatment cycle. Control group received standard IVF care.

Outcome measure Difference in variation due to SDM in the choice of SET or DET, both between and within hospitals.

Results There was large variation in the choice of SET or DET between hospitals in the control group. Lower variation between hospitals was observed in the group with SDM. Within most hospitals variation in the choice of SET or DET appeared to increase due to SDM. Variation particularly increased in hospitals where mainly DET was chosen in the control group.

Conclusions Although based on a limited dataset, our study gives a first insight that including patients' preferences through SDM results in less variation between hospitals, and indicates another pattern of variation within hospitals. Variation that results from patient preferences could be potentially named the informed patient rate. Our results provide the starting point for further research.

Trial registration ClinicalTrials.gov NCT00315029.

STRENGHTS AND LIMITATIONS OF THIS STUDY

- This study is the first to elaborate further upon, and explore the association between, SDM and medical practice variation.
- Data from an RCT are used, which enables a comparison to be drawn between a situation with, and one without, promoting SDM.
- A limitation is that we had only access to a limited dataset, and as such, we performed descriptive statistics to test our hypotheses.

INTRODUCTION

Considerable variation exists in medical treatment [1-4]. In a paternalistic model, the physician is the dominant actor deciding on this treatment [5 6]. This approach is widely practiced, embedded in the idea that physicians decide on treatment based on both medical science and what is best for an individual patient [3], i.e. the belief "the doctor knows best". As such, physicians' professional judgements rather than patients' preferences often determine which treatment a patient receives [3,p9 7]. As a result, variation is found to be related to physicians, rather than to patients [3 8 9]. In explaining variation, research therefore focuses on the role of physicians, while patients' influence receives little attention [10 11]. Research showed that variation, among others, can be explained by differences in opinions on, or enthusiasm for, certain procedures between individual physicians, and by differences in constraints and social influences for groups of physicians, for instance between hospitals [9 11-14]. In the past decades, however, the paternalistic model has been questioned. Also, the position of patients in health care has significantly altered. On an individual level they are supposed to take up an active role in their health [15], and are expected to be involved in decisions about their health [16]. There is, thus, an increased emphasis on including patients and their preferences in medical decision-making [17 18]. Since medical decision-making is decisive for variation in medical practice, it is questioned whether patients can still be ignored in theories about variation. Providing care that is respectful of, and responsive to, an individual patient's preferences, so-called patient-centredness, is regarded as of primary importance to health care alongside dimensions such as being safe, effective, timely, efficient, and equitable [19].

Medical decisions regarding treatment may change through the inclusion of patients' preferences as these preferences may deviate from physicians' professional judgements [20]. Including patients' preferences may result in different treatment choices, and patterns of variation. It has been hypothesized that patient involvement may reduce variation in medical practice, because research shows that patients, through a combination of education and participation, were less ready to accept certain procedures [21]. This also assumes that physicians are more diverse in their preferences than patients despite the fact that they have a shared training and socialisation that has no parallel among patients [21]. One specific approach to patient involvement is shared decision-making (SDM) [22 23]. SDM is especially important in case of preference sensitive care, i.e. when there is more than one clinically appropriate treatment option. SDM is defined as an approach where physicians and patients take decisions together using the best available evidence. Patients are helped to make informed choices by considering the options and the likely benefits and disadvantages of each option [24 25]. This is important as

informed patients often prefer other treatments than their physician [3]. Research showed that, in general, informed patients prefer less invasive treatment options [3 26]. For example, a study of Deyo et al. (2000) showed that patients with herniated disks who watched a video program chose less surgery [27]. On the other hand, variation exists between physicians, since some of them prefer invasive treatments and others conservative treatments [28]. As such, it has been suggested that SDM – as a case of patient involvement – reduces variation [23 29-31]. However, no clear evidence about the association between SDM and variation is available yet [11]. There is no research which has identified exactly how or why SDM might reduce variation. Therefore, this study further elaborates upon the mechanisms that may explain why SDM reduces practice variation. In addition, we aim to perform a first, explorative, analysis on a limited dataset to examine the hypothesis that SDM reduces medical practice variation. Hereby, we make use of a clear example of a decision which depends on patients' preferences, the choice of either a single embryo transfer (SET) or double embryo transfer (DET) after in vitro fertilisation (IVF) [32].

SET prevents a multiple pregnancy with associated higher risks. DET results in higher live birth rates per treatment cycle [33 34] [see Box 1 for more information]. The percentage of SET varies considerably between countries [35-37]. For example, rates of SET ranged from 8.7 per cent in Moldova to 70.7 per cent in Sweden [37]. Likewise, major differences exist in how this complex decision is taken. These differences exist between countries, and between hospitals within the same country. In some hospitals the decision is based solely on clinical parameters, while in other hospitals patients are fully involved in the decision and physicians act as advisor [38]. If the physician decides for SET or DET, this decision is mainly based on physicians' professional judgements. As such, variation is likely. When patients are involved, decisions may differ; informed patients often prefer less invasive treatments [3 26]. The data from the RCT analysed in this study, in which a strategy for SDM was used as intervention, showed that educated couples, who understood the risks of twin pregnancies, were more inclined to choose SET. This is compared to couples receiving standard care [32].

We examine how a strategy to promote SDM is associated with variation in choosing SET or DET both **between** and **within** hospitals. We hypothesise that SDM is associated with less variation **between** hospitals, since we expect that, due to SDM, SET is chosen more often both in hospitals where physicians already preferred SET and in hospitals where physicians preferred DET, since educated couples prefer this (H1). We also hypothesise that if DET is mainly preferred **within** a hospital, and there is thus hardly any variation, then SDM is expected to

increase variation, because SET will be chosen more often due to SDM (H2). Whereas, if SET is mainly preferred **within** a hospital, and there is thus hardly any variation, then SDM is expected not to change variation, since SET is still preferred due to SDM (H3). Furthermore, we hypothesise that if DET and SET are both chosen **within** a hospital, and there is thus large variation, SDM is expected to decrease variation, because, due to SDM, SET is likely to be chosen more often than DET (H4).

METHODS

Description of the data

Data for this research were obtained from the RCT by Van Peperstraten et al. (2010) [32]. The choice for SET or DET should ideally be decided in a SDM process by an educated and empowered couple. In the RCT study of Van Peperstraten et al. (2010) a multifaceted strategy was used to promote SDM. To promote SDM, Van Peperstraten et al. developed a decision aid (DA) [see Box 2 for more information]. DAs are standardized, evidence-based tools intended to promote SDM [23]. Besides the evidence-based DA, this strategy consisted of support of an IVF nurse and reimbursement for an additional cycle of IVF for couples for whom the choice of SET caused a reduced chance of pregnancy [32]. In the Netherlands up to three IVF cycles are covered by the basic (but extensive) health insurance. The content of the DA and the reimbursement offer were discussed in person with a trained IVF nurse. All three elements of the strategy were provided before the counseling session that was part of standard care [32]. The control group received standard IVF care, including a session discussing the choice of SET or DET. Next to this standard care, the intervention group received the multifaceted empowerment strategy [32]. In the original RCT study, participating women completed three questionnaires (at inclusion, after intervention (but before starting treatment) and five weeks after embryo transfer) to measure decision-making outcomes and knowledge. Results showed that the proportion of couples in the intervention group who wanted to decide for themselves on the number of embryos transferred increased, while this percentage remained the same in the control group ($p<0.001$). Levels of both experienced knowledge ($p=0.001$) and actual knowledge ($p<0.001$) where higher in the intervention group compared to the control group [32]. For further detailed information see Van Peperstraten et al. (2010) [32].

Before the study, in 2005, 39% of the couples underwent SET after the first cycle [39]. The RCT was performed in five hospitals in the Netherlands. It included couples on the IVF waiting list between November 2006 and July 2007. The follow-up was continued until December 2008. Couples of women under 40 were included if they were on the waiting list for their first IVF cycle ever or a first cycle after a previous successful IVF. Couples were

excluded if SET was mandatory due to a strict medical indication. Written informed consent was provided by the couples before participation [32].

Selection of the data

In total, 308 couples at the beginning of their first IVF cycle were included in the intention to treat analysis (ITT) of Van Peperstraten et al. (2010) [32]. In all five hospitals approximately half of the couples received standard care, while the other half received the intervention. In this study, only couples that had the opportunity to choose between SET and DET were included. We, therefore, omitted from the 308 couples included in the ITT all couples: 1) where the woman was pregnant before starting IVF (N=20); 2) that never started IVF (N=13); and 3) that had none or just one embryo available (N=39 respectively N=14). Our sample included 222 couples, 113 in the control group and 109 in the intervention group, respectively. The outcome measure used in this study was the choice of either SET or DET. The data on this outcome were collected by Van Peperstraten et al. (2010) from local IVF registries [32]. Other variables included were whether a couple was involved in the intervention group or in the control group, and the hospital in which a couple was treated. In addition, we included four variables that are of medical relevance and might, therefore, affect the choice of SET or DET, and thus practice variation. For example, the older the woman is, the less likely she will become pregnant and the more likely she will have twins. The four variables included were: 1) the age of the woman (in years); 2) the duration of infertility (in years); 3) the presence of a good quality embryo (yes/no); and 4) any previous pregnancies (yes/no). Data on the presence of a good quality embryo were collected by Van Peperstraten et al. (2010) from local IVF registries [32]. Data for the other three variables were collected through a patient questionnaire which couples received when included to the study [32]. The woman's age and duration of infertility were calculated in this study on 31 December 2008 (= end of follow-up).

Statistical analyses of the data

We examined whether the control and intervention group were comparable for the characteristics included by performing descriptive statistics, and chi^2 tests (categorical variables) and t-tests (continuous variables) ($p < 0.05$). We then examined whether the five hospitals included did significantly differ with respect to the four variables that are of medical relevance. If there were differences between the five hospitals then we had to take these into account throughout the rest of our analyses, since these may have an impact on the choice of SET or DET. We performed descriptive statistics per hospital for the four variables. By chi^2 tests (categorical variables) and one-

way analyses of variance (ANOVA) (continuous variables), we tested if there were significant differences between the five hospitals for the woman's age, the duration of infertility, the presence of a good quality embryo, and for previous pregnancies ($p < 0.05$). If a significant difference was found between the hospitals for one of the aforementioned variables, we then performed an additional analysis to examine if there was an association between that variable and the outcome measure.

We then calculated, for each hospital, the percentage of couples that chose SET or DET, both in the control and in the intervention groups. We examined this in order to confirm that educated couples are inclined to choose SET. Next, we examined the variation **between** the hospitals. We first calculated for each hospital the percentage of SET in the control group and then in the intervention group. Thereafter, we calculated the range of SET percentages for the control groups and the intervention groups. A smaller range or difference between the highest and the lowest percentage of SET, implies less variation. Thus, if a strategy to promote SDM is associated with less variation between hospitals, then the range of SET percentages for the intervention group is smaller than that for the control group.

We now examined the variation **within** the hospitals by looking at the differences in variation between the control and the intervention group in each hospital. We calculated for each hospital the absolute difference between SET and DET in the control as well as in the intervention group. For example, if 40% chose SET and 60% DET, the absolute difference is 20. There is no variation if the proportion of SET to DET is 0% compared to 100%, or vice versa. Thus, there is no variation if the absolute difference is 100. On the other hand, the most variation is observed if the proportion of SET to DET is 50% to 50%. Thus, there is an absolute difference of 0. We can thus create a scale ranging from 0 to 100, where a score closer to 100 means less variation. We compared the scores of the control and the intervention group for each hospital. If the score in a hospital is higher in the intervention group than in the control group, and thus closer to 100 (= no variation), then a strategy to promote SDM is associated with less variation within that hospital. Complementary to the descriptive statistics, we performed a multilevel analysis (MLA) in MLwiN to examine variation between hospitals. A MLA takes into account the nested structure of the data as well as the differences in the number of patients per hospital. All statistical analyses were carried out using STATA, version 13.1.

RESULTS

Characteristics of the couples included

A description of the 222 couples included is given in Table 1. The number of couples included ranged from twelve couples in hospital five to 153 couples in hospital one (see Table 1). The control and intervention group did not differ significantly with respect to the characteristics included. Furthermore, no significant differences were observed between the five hospitals for the mean duration of infertility (p=0.256), the presence of a good quality embryo (p=0.406), and for previous pregnancies (p=0.403) (see Table 2). ANOVA showed a significant difference (p=0.032) for the variable, woman's age. An additional t-test showed no difference between woman's age and the choice of SET or DET (p=0.346). Thus, we decided not to include these four variables throughout the rest of the analyses.

Choice of SET

Table 3 shows the numbers and percentages of SET and DET for both the control and intervention groups, in total and per hospital. In total, 52% of the couples included in the intervention group chose SET, in comparison with 39% of the couples in the control group (p=0.046). To be more specific, in four of the five hospitals, couples in the intervention groups more often chose SET than DET. In hospital four, however, couples in the intervention group more often chose DET than SET (80% vs. 20%). Although in hospital four couples in the intervention group more often chose DET, they more often chose SET (20%) than couples in the control group (0%).

Variation between hospitals

The range of SET in both the control and intervention groups can also be observed in Table 3. The percentages of SET in the control groups ranged from 0.0% to 85.7%, while the percentages of SET in the intervention groups ranged from 20.0% to 87.5%. Therefore the range of SET is smaller in the intervention group than in the control group, which is an indication that a strategy to promote SDM reduced variation in the choice of SET or DET between hospitals. The MLA also indicated that the variation between hospitals was lower in the intervention group than in the control group. However, the difference was not significant.

Variation within hospitals

Figure 1 shows the differences in variation within hospitals by illustrating, per hospital, the absolute difference between SET and DET in the control group (standard care) and the intervention group (strategy to promote SDM). In one hospital (number 2) the absolute difference in the control group and the intervention group is the same. This means that the variation within hospital 2 is the same with or without a strategy to promote SDM. Within

hospital 3, the strategy to promote SDM appears to be associated with less variation, since the absolute difference in the control group is lower than in the intervention group (14 and 75, respectively). On the other hand, within the other three hospitals, numbers 1, 4 and 5, the strategy to promote SDM appears to be associated with more variation. Within these three hospitals the absolute difference in the control group is higher than in the intervention group (see Figure 1). Therefore, within some of the hospitals included, a strategy to promote SDM appears to be associated with more variation, while within other hospitals a strategy to promote SDM appears to be associated with less or the same level of variation.

DISCUSSION AND CONCLUSIONS

Principal findings

This study further elaborated upon and explored the association between SDM and variation in the choice of SET or DET both between and within hospitals. There was large variation in the choice of SET or DET between hospitals in the control group. Lower variation between hospitals was observed in the group with a strategy to promote SDM. Furthermore, we observed that within most hospitals the variation in the choice of SET or DET appeared to increase due to a strategy to promote SDM. This was particularly in hospitals where mainly DET was chosen in the control group.

What this study adds

Literature suggests that SDM reduces variation [23 29-31]. There was however, up to now, no clear evidence about this association. This study is the first that explored this association based on a case concerning the choice of SET or DET after IVF. We noticed that a strategy to promote SDM reduces variation between hospitals (confirming H1), while the variation within most hospitals appears to increase. The hypothesis in literature that SDM reduces variation is based on the observation that informed patients more often prefer less invasive treatments [3 26]. We found that in most hospitals couples in the intervention group more often chose SET. Although this does not imply that there will be less variation, since our results indicate that variation within most hospitals increased. This is because the level of variation without SDM differed between hospitals. For example, in some of the hospitals included mainly DET was preferred and there was thus almost no variation. Due to the strategy to promote SDM, however, SET was chosen more often, and thus the variation increased within such hospitals, since now both SET and DET are chosen (confirming H2). In one hospital SET was mainly chosen in the control group, and we therefore expected no change in variation, since we expected SET still preferred due to

the strategy to promote SDM (H3). However, we observed an increase in variation in that hospital, rejecting H3. In the two hospitals with the largest variation in the control group the variation decreased or remained equal, confirming H4. A subsequent implication is that an overall decrease in variation between hospitals, provides no indication about the change in variation within an individual hospital. Although based on a limited dataset, this study gives a first insight that SDM results in less variation between hospitals while suggesting another pattern of variation within hospitals.

Further research

This research focused on just one decision-making situation in obstetrics, and had only access to a limited dataset. The results, therefore, have to be interpreted with caution and further research is necessary both to underpin our results and to examine questions that remain unanswered. Nevertheless, our study provides a starting point for further empirical research within this area. For many medical problems no absolute best treatment option is available and so there are significant trade-offs among the available options [40 41]. We expect, however, that our results apply generally to medical problems with no absolute best treatment option. Decisions concerning such problems are defined as preference sensitive, since they depend on considerations of the benefits, disadvantages, and uncertainties of each treatment. For example some patients will prefer to accept a small risk of death in order to attempt to improve their function, while others will not [23 41]. Therefore, the best decisions cannot be made without including patients' preferences [40 41]. Well-known examples include chronic back pain, early-stage breast cancer and prostate cancer. For examples such as these it is believed that variation will change as a result of SDM. Future research has to confirm this by making use of data from multicentre RCT studies that applied intervention strategies (like a DA) to increase SDM in a specific consultation. Such RCT studies have been carried out [26], but have focused, comparable to the study we used, on outcome measures other than variation. Therefore, we decided in this study to perform a secondary analysis. Any possible multicentre studies should include a control and an intervention group which could thus measure actual treatment choices with and without SDM. This would allow researchers to examine whether SDM changes the pattern of variation, by, for example, using the same method as we did. Another possibility for further research is to conduct a new multicentre RCT study specifically aimed at analysing the relationship between SDM and medical practice variation.

Our results show that a strategy to promote SDM results in less variation between hospitals and indicates another pattern of variation within hospitals, confirming our hypotheses. These results appear to show that the decisions

made by informed patients have a pattern too. Choices made by informed patients appear to have a rate which deviates from baseline rates, irrespective of whether those are "low" or "high". This could be potentially named the informed patient rate. However, it can be questioned whether the rates we observed are indeed the informed patient rate, that is the results of what the couples want. It is possible that not all patients were able, or preferred, to take a shared decision about the choice of SET or DET. Nevertheless, results of the original RCT study show that levels of both experienced knowledge ($p=0.001$) and actual knowledge ($p<0.001$) where higher in the intervention group compared to the control group [32].Further research has to examine whether the actual choice was indeed the patients' preference and whether there are differences between groups of patients in this. In addition, a different pattern of variation due to SDM might be a positive indication for the quality of care. Good health care requires, among others, providing care that is respectful of, and responsive to, individual patients' preferences [19]. This is particularly true for preference sensitive decisions, since these decisions depend on patients' preferences regarding the benefits, disadvantages and uncertainties of each treatment. Further research has to examine whether SDM results in better quality of care for preference sensitive decisions.

The broad context of this study is about the influence of patients and their preferences on variation in medical practice. SDM is one option for including patients' preferences in medical decision-making. There are other options through which patients can express their preferences, and thus to influence the pattern of variation. For example, patients differ in how much pressure they are able to put on physicians [13 42]. They differ in their ability to take part in discussions over treatment with their physicians. Some patients are expected to be able to ask their physician for another treatment than, for example, the treatment that is recommended in a guideline or the one that is preferred by the physician [13 43]. If this is the case, then patients' preferences appear to influence the treatment chosen, and thus the variation. Further research is recommended into these situations.

The strengths and weaknesses of the study

A strength of our study is that we are the first to elaborate further upon and explore the association between SDM and variation in medical treatment. We examined this association to get insight into whether including patients' preferences through a strategy to promote SDM results in another pattern of variation in medical treatment. Another strength is the use of data from a RCT. We had the opportunity to compare the variation in the choice of SET or DET with, and without, SDM. It might be possible that the choice of SET or DET in the control group is influenced by physicians, since they treated both couples in the intervention and in the control group. Ideally, data

would have been available about the percentage of SET and DET before the RCT, allowing us to compare the intervention and control group with these percentages. Though another study showed that in 2005, before the RCT, 39% of twin prone couples in two Dutch hospitals chose SET [39], which is comparable to the percentage of SET in the control group. It seems plausible to use the control group as the situation before SDM. We performed descriptive statistics to analyse our data, because of the low numbers of couples included in the hospitals. We have taken into account the nested structure of the data by performing our analyses per hospital. However, we did not take into account the differences in the number of patients per hospital. We therefore also performed an MLA to examine the variation between hospitals. The MLA supported the results of the descriptive statistics, however, the difference was not significant. From the dataset it was also known only in which hospital a couple was treated, but not by which physician within that hospital. However, only one or two physicians per hospital treated all couples in that hospital so we do not expect that this will affect our conclusions. Further research should ideally be performed with a larger dataset, preferably using multilevel analysis in order to test the hypothesis of this study. This would acknowledge that patients are nested hierarchically within physicians and physicians within hospitals. In addition, further research has to include in the analyses socio-demographics that might have an influence on the treatment decision. A final limitation might be that the intervention consisted of different elements, and thus it is difficult to assess separately the effects of these elements. Despite the fact that at the end of the follow-up period only 4% of the couples qualified for reimbursement of a fourth cycle, reimbursement might have played a role in the decision [32]. However, a follow-up study showed that – compared to the other elements – the reimbursement offer was rated least important by the couples in choice for SET or DET [44].

Conclusions

This study was the first to elaborate further upon and explore the relationship between including patients' preferences in medical decision-making and practice variation. Although based on a limited dataset, our study gives a first insight that including patients' preferences through SDM results in less variation between hospitals, and indicates another pattern of variation within hospitals. The variation that results from patient preferences could be potentially named the informed patient rate. The results of this study provide the starting point for further empirical research within this area.

Acknowledgements We thank Peter Spreeuwenberg (NIVEL) for advising on and performing the MLA.

Authorship AB was involved in the design of the study, performed the statistical analyses, and drafted the manuscript. JDJ and LVD were involved in the design of the study, assisted in interpreting the results, assisted in drafting the manuscript, and revised the manuscript. PG was involved in the design of the study, and revised the manuscript. AVP was the primary author of the original RCT and was therefore responsible for the data collection, and revised the manuscript. All authors have read and approved the final manuscript.

Conflict of interest All authors have completed the unified competing interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare (1) all no financial support for the work submitted from anyone other than their employer; (2) AB, PG, AVP and JDJ no financial relationships with commercial entities that might have an interest in the work submitted. LVD received unrestricted grants from Bristol-Myers Squibb, Pfizer and Astra Zeneca for studies not related to this study; (3) all no spouses, partners, or children with relationships with commercial entities that might have an interest in the work submitted; and (4) all no non-financial interests that may be relevant to the submitted work.

Source of funding The RCT study was funded by the Netherlands Organisation for Health Research and Development (grant no 945-16-105). For this secondary analysis of the RCT no additional funding was obtained.

Ethics approval The RCT study was approved by the regional ethics committee for medical research.

Data sharing statement The dataset of the RCT study is available upon request from AVP.

Transparency All authors are a guarantor and affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no imported aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

REFERENCES

1. Paul-Shaheen P, Clark JD, Williams D. Small area analysis: a review and analysis of the North American literature. J Health Polit Policy Law 1987;**12**(4):741-809.

2. Wennberg JE. Understanding geographic variations in health care delivery. NEJM 1999;**340**(1):52-3.

3. Wennberg JE. *Tracking medicine: a researcher's quest to understand health care*: Oxford University Press, 2010.

4. Corallo AN, Croxford R, Goodman DC, et al. A systematic review of medical practice variation in OECD countries. Health Policy 2014;**114**(1):5-14.

5. Emanuel EJ, Emanuel LL. Four models of the physician-patient relationship. JAMA 1992;**267**(16):2221-6.

6. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med 1997;**44**(5):681-92.

7. Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. BMJ 2002;**325**(7370):961.

8. McPherson K. Why do variations occur? In: Andersen TF, Mooney G, eds. The challenges of medical practice variations. London: The Macmillan Press, 1990:16-35.

9. Wennberg JE, Gittelsohn A. Health care delivery in Maine I: patterns of use of common surgical procedures. J Maine Med Assoc 1975;**66**(5):123-30, 49.

10. Mercuri M, Gafni A. Medical practice variations: what the literature tells us (or does not) about what are warranted and unwarranted variations. J Eval Clin Pract 2011;**17**(4):671-7.

11. De Jong JD, Groenewegen PP, Westert GP. Sociological Model for Understanding Medical Practice Variations. In: Sobolev B (Series ed) JA, Stukel T (eds). , ed. Health Services Research Series Medical Practice Variations. New York: Springer, 2015.

12. Chassin MR. Explaining geographic variations. The enthusiasm hypothesis. Med Care 1993;**31**(5 Suppl):Ys37-44.

13. De Jong JD. Explaining medical practice variation. Social organization and institutional mechanisms. NIVEL, 2008.

14. Westert GP, Groenewegen PP. Medical practice variations: changing the theoretical approach. Scand J Public Health 1999;**27**(3):173-80.

15. van de Bovenkamp HM, Trappenburg MJ, Grit KJ. Patient participation in collective healthcare decision making: the Dutch model. Health Expect 2010;**13**(1):73-85.

16. Guadagnoli E, Ward P. Patient participation in decision-making. Soc Sci Med 1998;**47**(3):329-39.

17. Frosch DL, Kaplan RM. Shared decision making in clinical medicine: past research and future directions. Am J Prev Med 1999;**17**(4):285-94.

18. Fraenkel L. Incorporating patients' preferences into medical decision making. Med Care Res Rev 2013;**70**(1 Suppl):80s-93s.

19. Institute of Medicine. *Crossing the quality chasm: A new health system for the 21st century*: National Academies Press, 2001.

20. Wennberg JE. *Variation in Use of Medicare Services among Regions and Selected Academic Medical Centers: Is More Better?*: Commonwealth Fund, 2005.

21. Greer AL, Goodwin JS, Freeman JL, et al. Bringing the patient back in. Guidelines, practice variations, and the social context of medical practice. Int J Technol Assess Health Care 2002;**18**(4):747-61.

22. Charles C, Gafni A, Whelan T. Decision-making in the physician-patient encounter: revisiting the shared treatment decision-making model. Soc Sci Med 1999;**49**(5):651-61.

23. O'Connor AM, Llewellyn-Thomas HA, Flood AB. Modifying unwarranted variations in health care: shared decision making using patient decision aids. Health Affairs 2004:Var63-72.

24. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012;**27**(10):1361-7.

25. Elwyn G, Laitner S, Coulter A, et al. Implementing shared decision making in the NHS. BMJ 2010;**341**:c5146.

26. Stacey D, Bennett CL, Barry MJ, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2011(10):Cd001431.

27. Deyo RA, Cherkin DC, Weinstein J, et al. Involving patients in clinical decisions: impact of an interactive video program on use of back surgery. Med Care 2000;**38**(9):959-69.

28. McClure W. Toward development and application of a qualitative theory of hospital utilization. Inquiry 1982;**19**(2):117-35.

29. O'Connor AM, Wennberg JE, Legare F, et al. Toward the 'tipping point': decision aids and informed patient choice. Health Affairs 2007;**26**(3):716-25.

30. Stiggelbout AM, Van der Weijden T, De Wit MP, et al. Shared decision making: really putting patients at the centre of healthcare. BMJ 2012;**344**:e256.

31. Weinstein JN, Bronner KK, Morgan TS, et al. Trends and geographic variations in major surgery for degenerative diseases of the hip, knee, and spine. Health Affairs 2004:Var81-9.

32. van Peperstraten A, Nelen W, Grol R, et al. The effect of a multifaceted empowerment strategy on decision making about the number of embryos transferred in in vitro fertilisation: randomised controlled trial. BMJ 2010;**341**:c2501.

33. Pandian Z, Marjoribanks J, Ozturk O, et al. Number of embryos for transfer following in vitro fertilisation or intra-cytoplasmic sperm injection. Cochrane Database Syst Rev 2013;**7**:Cd003416.

34. Pinborg A. IVF/ICSI twin pregnancies: risks and prevention. Hum Reprod Update 2005;**11**(6):575-93.

35. Andersen AN, Goossens V, Ferraretti AP, et al. Assisted reproductive technology in Europe, 2004: results generated from European registers by ESHRE. Hum Reprod 2008;**23**(4):756-71.

36. de Mouzon J, Goossens V, Bhattacharya S, et al. Assisted reproductive technology in Europe, 2006: results generated from European registers by ESHRE. Hum Reprod 2010;**25**(8):1851-62.

37. Ferraretti AP, Goossens V, Kupka M, et al. Assisted reproductive technology in Europe, 2009: results generated from European registers by ESHRE. Hum Reprod 2013;**28**(9):2318-31.

38. van Peperstraten A. Implementation of single embryo transfer: a patient directed strategy. Radboud Universiteit Nijmegen, 2009.

39. Van Peperstraten AM, Kreuwel IA, Hermens RP, et al. Determinants of the choice for single or double embryo transfer in twin prone couples. Acta Obstet Gynecol Scand 2008;**87**(2):226-31.

40. Barry MJ. Shared decision making: informing and involving patients to do the right thing in health care. J Ambul Care Manage 2012;**35**(2):90-8.

41. Dartmouth Atlas. A Dartmouth Atlas Project Topic Brief Preference-Sensitive Care. Lebanon: The Dartmouth Atlas of Health Care, 2007.

42. Westert GP, Groenewegen PP, Lindenberg SM. De invloed van ziekenhuispatiënten op de duur van de hospitalisatie. TSG 1991;**69**:184-94.

43. Hajjaj FM, Salek MS, Basra MK, et al. Non-clinical influences on clinical decision-making: a major challenge to evidence-based practice. J R Soc Med 2010;**103**(5):178-87.

44. Kreuwel IA, van Peperstraten AM, Hulscher ME, et al. Evaluation of an effective multifaceted implementation strategy for elective single-embryo transfer after in vitro fertilization. Hum Reprod 2013;**28**(2):336-42.

45. No authors listed. Prevention of twin pregnancies after IVF/ICSI by single embryo transfer. ESHRE Campus Course Report. Hum Reprod 2001;**16**(4):790-800.

46. Wolner-Hanssen P, Rydhstroem H. Cost-effectiveness analysis of in-vitro fertilization: estimated costs per successful pregnancy after transfer of one or two embryos. Hum Reprod 1998;**13**(1):88-94.

47. Lukassen HG, Schonbeck Y, Adang EM, et al. Cost analysis of singleton versus twin pregnancies after in vitro fertilization. Fertil Steril 2004;**81**(5):1240-6.

48. van Peperstraten AM, Hermens RP, Nelen WL, et al. Deciding how many embryos to transfer after in vitro fertilisation: development and pilot test of a decision aid. Patient Educ Couns 2010;**78**(1):124-9.

49. Elwyn G, O'Connor A, Stacey D, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ 2006;**333**(7565):417.

BOXES

Box 1: The choice of single or double embryo transfer: A complex decision-making problem

The choice of a single or double embryo transfer after IVF is a complex decision-making problem because of the need to find a balance between the risk of complications of multiple birth and the best chance of pregnancy [32]. Some sub-fertile couples and professionals regard twin pregnancies as a success, however, they could also be considered as a side effect or even a complication [45]. Twin pregnancies are associated with higher morbidity and mortality rates for both mother and child compared to singleton pregnancies [34]. Moreover, complications of twin pregnancies cause substantial use of medical budgets [46 47]. Subsequently, twin pregnancies are increasingly regarded as undesirable. To prevent twin pregnancies professionals and couples could choose single embryo transfer (SET) instead of double embryo transfer (DET) [34 45]. However, this may be disadvantageous, since it could result in a lower pregnancy rate per IVF cycle [33]. The choice of SET or DET is ideally decided through SDM [48].

Box 2: The choice of single or double embryo transfer: An evidence-based decision aid

Van Peperstraten et al. (2010) developed and tested the evidence-based decision aid (DA) for deciding how many embryos to transfer during IVF [48]. The DA was developed according to the checklist of the International Patient Decision Aids Standards Collaboration, which consists of 50 items divided between three domains, content, development, and effectiveness [48 49]. The purpose of the DA is to give couples all the information needed to make the choice to transfer one or two embryos and to relate the information to their own personal situation. The DA consists of three chapters: 1) information about the chances of a single pregnancy or a twin pregnancy; 2) information about the risks of twin pregnancies; and 3) an explanation of the available options and an action plan [48]. The DA is available in English at: www.umcn.nl/ivfda-en.

TABLES

Table 1. Characteristics of the couples included (N = 222). Values are numbers unless otherwise stated.

Characteristics	Control group (N = 113)	Intervention group (N = 109)	Total (N = 222)	p-value[c]
Hospital				
hospital 1	79	74	153	N.A.
hospital 2	7	7	14	
hospital 3	7	8	15	
hospital 4	13	15	28	
hospital 5	7	5	12	
Mean (sd) age of woman (years)[a]	33.9 (3.85) (range 21-41 years)	33.5 (3.88) (range 25-41 years)	33.7 (3.86) (range 21-41 years)	0.475
Mean (sd) duration of infertility (years)[a]	4.03 (2.08) (range 1-13 years) (N=101)	3.94 (1.91) (range 1-12 years) (N=98)	3.98 (2.00) (range 1-13 years) (N=199)	0.749
Presence of a good quality embryo				
no	41	28	69	0.088
yes	72	81	153	
Previous pregnancies[b]	(N=113)	(N=108)	(N=221)	
no	63	63	126	0.698
yes	50	45	95	

[a] Calculated on December 31, 2008 based on information filled out in patients' questionnaires. As a result, we have a higher mean for age and duration of infertility than Van Peperstraten et al., BMJ, 2010.

[b] Based on the question: "Have you ever been pregnant?"

[c] $p < 0.05$ is significant

Table 2. Characteristics of the couples included per hospital.

Characteristics	Total (N = 222)	p-value[c]
Mean (sd) age of woman (years)[a]	33.7 (3.86) (21-41)	
hospital 1 (N = 153)	33.8 (3.63) (21-41)	**0.032**
hospital 2 (N = 14)	30.9 (4.54) (25-39)	
hospital 3 (N = 15)	34.6 (3.96) (28-40)	
hospital 4 (N = 28)	33.4 (4.53) (25-41)	
hospital 5 (N = 12)	35.2 (2.86) (30-38)	
Mean (sd) duration of infertility (years)[a]	4.0 (2.00) (1-13)	
hospital 1 (N = 139)	4.1 (2.20) (1-13)	0.256
hospital 2 (N = 11)	4.0 (1.26) (2-6)	
hospital 3 (N = 11)	4.3 (1.85) (2-8)	
hospital 4 (N = 26)	3.2 (1.22) (1-6)	
hospital 5 (N = 12)	4.0 (1.13) (2-6)	
Presence of a good quality embryo (% yes)	68.9%	
hospital 1 (N = 153)	72.6%	0.406
hospital 2 (N = 14)	64.3%	
hospital 3 (N = 15)	66.7%	
hospital 4 (N = 28)	60.7%	
hospital 5 (N = 12)	50.0%	
Previous pregnancies[b] (% yes)	43.0%	
hospital 1 (N = 153)	45.8%	0.403
hospital 2 (N = 14)	21.4%	
hospital 3 (N = 14)	50.0%	
hospital 4 (N = 28)	39.3%	
hospital 5 (N = 12)	33.3%	

[a] Calculated on December 31, 2008 based on information filled out in patients' questionnaires. As a result, we have a higher mean for age and duration of infertility than Van Peperstraten et al., BMJ, 2010.

[b] Based on the question: "Have you ever been pregnant?"

[c] $p < 0.05$ is significant

Table 3. The choice of SET or DET total group, and per hospital. Values are numbers (%) unless otherwise stated.

	Control group	Intervention group	p-value[a]
Total			
SET	44 (38.9%)	57 (52.3%)	**0.046**
DET	69 (61.1%)	52 (47.7%)	
Hospital 1			
SET	31 (39.2%)	40 (54.1%)	0.066
DET	48 (60.8%)	34 (46.0%)	
Hospital 2			
SET	3 (42.9%)	4 (57.1%)	0.593
DET	4 (57.1%)	3 (42.9%)	
Hospital 3			
SET	4 (57.1%)	7 (87.5%)	0.185
DET	3 (42.9%)	1 (12.5%)	
Hospital 4			
SET	0 (0.0%)	3 (20.0%)	0.088
DET	13 (100.0%)	12 (80.0%)	
Hospital 5			
SET	6 (85.7%)	3 (60.0%)	0.310
DET	1 (14.3%)	2 (40.0%)	

[a] $p < 0.05$ is significant

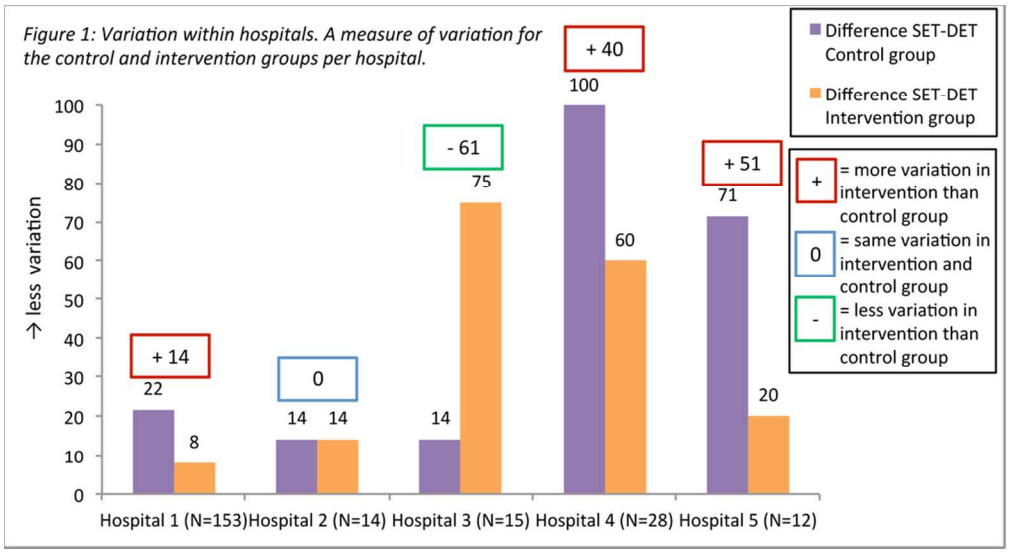


Figure 1: Variation within hospitals. A measure of variation for the control and intervention groups per hospital.
94x51mm (300 x 300 DPI)

CONSORT 2010 checklist of information to include when reporting a randomised trial*

We performed a secondary analysis of a RCT. The original RCT was published in the BMJ (2010) as "The effect of a multifaceted empowerment strategy on decision making about the number of embryos transferred in in vitro fertilisation: randomised controlled trial" by Van Peperstraten et al. We included this article as supplement to our submission.

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	See Van Peperstraten et al. (2010)
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	Page 2 of this submission
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	Page 4-5 of this submission
	2b	Specific objectives or hypotheses	Page 4-5 of this submission
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	See Van Peperstraten et al. (2010)
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	NA
Participants	4a	Eligibility criteria for participants	See Van Peperstraten et al. (2010) and Page 6 of this submission
	4b	Settings and locations where the data were collected	See Van Peperstraten et al. (2010) and Page 6 of this submission
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	See Van Peperstraten et al. (2010) and Page 6 of this submission
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	Page 6-7 of this submission

	6b	Any changes to trial outcomes after the trial commenced, with reasons	NA
Sample size	7a	How sample size was determined	See Van Peperstraten et al. (2010) en Page 6 of this submission
	7b	When applicable, explanation of any interim analyses and stopping guidelines	NA
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	See Van Peperstraten et al. (2010)
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	See Van Peperstraten et al. (2010)
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	See Van Peperstraten et al. (2010)
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	See Van Peperstraten et al. (2010)
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	NA
	11b	If relevant, description of the similarity of interventions	NA
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	Page 7-8 of this submission
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	NA
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	Table 1 and 2 and 3 of this submission
	13b	For each group, losses and exclusions after randomisation, together with reasons	Page 6 of this submission
Recruitment	14a	Dates defining the periods of recruitment and follow-up	See Van Peperstraten et al. (2010) and Page 6 of this submission
	14b	Why the trial ended or was stopped	NA
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Table 1 and 2 of this submission
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Table 1 and 2 and 3 of this submission
Outcomes and	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its	Table 3 and Figure 1 and

estimation		precision (such as 95% confidence interval)	Page 8-9 of this submission
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	Table 3 and Figure 1
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	NA
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	NA
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	See Van Peperstraten et al. (2010) and Page 12 of this submission
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Page 10 of this submission
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	Page 10-12 of this submission
Other information			
Registration	23	Registration number and name of trial registry	See Van Peperstraten et al. (2010) and Page 2 of this submission
Protocol	24	Where the full trial protocol can be accessed, if available	NA
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	See Van Peperstraten et al. (2010) and Page 13 of this submission

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

BMJ Open

Does a strategy to promote shared decision-making reduce medical practice variation in the choice of either single or double embryo transfer after in vitro fertilisation? A secondary analysis of an randomised controlled trial.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010894.R2
Article Type:	Research
Date Submitted by the Author:	29-Mar-2016
Complete List of Authors:	Brabers, Anne; NIVEL Netherlands Institute for Health Services Research, van Dijk, Liset; NIVEL Netherlands institute for health services research, Groenewegen, Peter; NIVEL Netherlands institute for health services research; Utrecht University, Department of Sociology, Department of Human Geography van Peperstraten, Arno; Radboud University Medical Center, Department of Obstetrics and Gynaecology De Jong, Judith; NIVEL Netherlands institute for health services research,
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Health policy, Obstetrics and gynaecology
Keywords:	shared decision-making, medical practice variation, in vitro fertilisation, patients' preferences, patient involvement

Does a strategy to promote shared decision-making reduce medical practice variation in the choice of either single or double embryo transfer after in vitro fertilisation? A secondary analysis of an randomised controlled trial.

Anne E.M. Brabers MSc[1§], Liset van Dijk PhD[1], Peter P. Groenewegen PhD[12], Arno M. van Peperstraten PhD[3], Judith D. de Jong PhD[1]

[1]NIVEL, the Netherlands Institute for Health Services Research, PO Box 1568, 3500 BN Utrecht, the Netherlands

[2]Department of Sociology, Department of Human Geography, Utrecht University, PO Box 80125, 3508 TC Utrecht, the Netherlands

[3]Department of Obstetrics and Gynaecology, Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, the Netherlands

[§]Corresponding author

Contact address corresponding author: PO Box 1568, 3500 BN Utrecht, the Netherlands, T: +31 (0)30 27 29 700, F: +31 (0)30 27 29 729

Email addresses:

AB: a.brabers@nivel.nl

LVD: l.vandijk@nivel.nl

PG: p.groenewegen@nivel.nl

AVP: arnovp@outlook.com

JDJ: j.dejong@nivel.nl

Word count: 4,583

Key-words shared decision-making, medical practice variation, in vitro fertilisation, patients' preferences, patient involvement

ABSTRACT

Objectives The hypothesis that shared decision-making (SDM) reduces medical practice variation is increasingly common, but no evidence is available. We aimed to elaborate further on this, and to perform a first, exploratory, analysis to examine this hypothesis. This analysis, based on a limited dataset, examined how SDM is associated with variation in the choice of single embryo transfer (SET) or double embryo transfer (DET) after in vitro fertilisation (IVF). We examined variation between and within hospitals.

Design A secondary analysis of an RCT.

Setting Five hospitals in The Netherlands.

Participants 222 couples (woman aged <40) on a waiting list for a first IVF cycle, who could choose between SET and DET (i.e. ≥ 2 embryos available).

Intervention SDM via a multifaceted strategy aimed to empower couples in deciding how many embryos should be transferred. The strategy consisted of decision aid, support of IVF nurse, and the offer of reimbursement for an extra treatment cycle. Control group received standard IVF care.

Outcome measure Difference in variation due to SDM in the choice of SET or DET, both between and within hospitals.

Results There was large variation in the choice of SET or DET between hospitals in the control group. Lower variation between hospitals was observed in the group with SDM. Within most hospitals variation in the choice of SET or DET appeared to increase due to SDM. Variation particularly increased in hospitals where mainly DET was chosen in the control group.

Conclusions Although based on a limited dataset, our study gives a first insight that including patients' preferences through SDM results in less variation between hospitals, and indicates another pattern of variation within hospitals. Variation that results from patient preferences could be potentially named the informed patient rate. Our results provide the starting point for further research.

Trial registration ClinicalTrials.gov NCT00315029.

STRENGHTS AND LIMITATIONS OF THIS STUDY

- This study is the first to elaborate further upon, and explore the association between, SDM and medical practice variation.
- Data from an RCT are used, which enables a comparison to be drawn between a situation with, and one without, promoting SDM.
- A limitation is that we had only access to a limited dataset, and as such, we performed descriptive statistics to test our hypotheses.

INTRODUCTION

Considerable variation exists in medical treatment [1-4]. In a paternalistic model, the physician is the dominant actor deciding on this treatment [5 6]. This approach is widely practiced, embedded in the idea that physicians decide on treatment based on both medical science and what is best for an individual patient [3], i.e. the belief "the doctor knows best". As such, physicians' professional judgements rather than patients' preferences often determine which treatment a patient receives [3,p9 7]. As a result, variation is found to be related to physicians, rather than to patients [3 8 9]. In explaining variation, research therefore focuses on the role of physicians, while patients' influence receives little attention [10 11]. Research showed that variation, among others, can be explained by differences in opinions on, or enthusiasm for, certain procedures between individual physicians, and by differences in constraints and social influences for groups of physicians, for instance between hospitals [9 11-14]. In the past decades, however, the paternalistic model has been questioned. Also, the position of patients in health care has significantly altered. On an individual level they are supposed to take up an active role in their health [15], and are expected to be involved in decisions about their health [16]. There is, thus, an increased emphasis on including patients and their preferences in medical decision-making [17 18]. Since medical decision-making is decisive for variation in medical practice, it is questioned whether patients can still be ignored in theories about variation. Providing care that is respectful of, and responsive to, an individual patient's preferences, so-called patient-centredness, is regarded as of primary importance to health care alongside dimensions such as being safe, effective, timely, efficient, and equitable [19].

Medical decisions regarding treatment may change through the inclusion of patients' preferences as these preferences may deviate from physicians' professional judgements [20]. Including patients' preferences may result in different treatment choices, and patterns of variation. It has been hypothesized that patient involvement may reduce variation in medical practice, because research shows that patients, through a combination of education and participation, were less ready to accept certain procedures [21]. This also assumes that physicians are more diverse in their preferences than patients despite the fact that they have a shared training and socialisation that has no parallel among patients [21]. One specific approach to patient involvement is shared decision-making (SDM) [22 23]. SDM is especially important in case of preference sensitive care, i.e. when there is more than one clinically appropriate treatment option. SDM is defined as an approach where physicians and patients take decisions together using the best available evidence. Patients are helped to make informed choices by considering the options and the likely benefits and disadvantages of each option [24 25]. This is important as

informed patients often prefer other treatments than their physician [3]. Research showed that, in general, informed patients prefer less invasive treatment options [3 26]. For example, a study of Deyo et al. (2000) showed that patients with herniated disks who watched a video program chose less surgery [27]. On the other hand, variation exists between physicians, since some of them prefer invasive treatments and others conservative treatments [28]. As such, it has been suggested that SDM – as a case of patient involvement – reduces variation [23 29-31]. However, no clear evidence about the association between SDM and variation is available yet [11]. There is no research which has identified exactly how or why SDM might reduce variation. Therefore, this study further elaborates upon the mechanisms that may explain why SDM reduces practice variation. In addition, we aim to perform a first, explorative, analysis on a limited dataset to examine the hypothesis that SDM reduces medical practice variation. Hereby, we make use of a clear example of a decision which depends on patients' preferences, the choice of either a single embryo transfer (SET) or double embryo transfer (DET) after in vitro fertilisation (IVF) [32].

SET prevents a multiple pregnancy with associated higher risks. DET results in higher live birth rates per treatment cycle [33 34] [see Box 1 for more information]. The percentage of SET varies considerably between countries [35-37]. For example, rates of SET ranged from 8.7 per cent in Moldova to 70.7 per cent in Sweden [37]. Likewise, major differences exist in how this complex decision is taken. These differences exist between countries, and between hospitals within the same country. In some hospitals the decision is based solely on clinical parameters, while in other hospitals patients are fully involved in the decision and physicians act as advisor [38]. If the physician decides for SET or DET, this decision is mainly based on physicians' professional judgements. As such, variation is likely. When patients are involved, decisions may differ; informed patients often prefer less invasive treatments [3 26]. The data from the RCT analysed in this study, in which a strategy for SDM was used as intervention, showed that educated couples, who understood the risks of twin pregnancies, were more inclined to choose SET. This is compared to couples receiving standard care [32].

We examine how a strategy to promote SDM is associated with variation in choosing SET or DET both **between** and **within** hospitals. We hypothesise that SDM is associated with less variation **between** hospitals, since we expect that, due to SDM, SET is chosen more often both in hospitals where physicians already preferred SET and in hospitals where physicians preferred DET, since educated couples prefer this (H1). We also hypothesise that if DET is mainly preferred **within** a hospital, and there is thus hardly any variation, then SDM is expected to

increase variation, because SET will be chosen more often due to SDM (H2). Whereas, if SET is mainly preferred **within** a hospital, and there is thus hardly any variation, then SDM is expected not to change variation, since SET is still preferred due to SDM (H3). Furthermore, we hypothesise that if DET and SET are both chosen **within** a hospital, and there is thus large variation, SDM is expected to decrease variation, because, due to SDM, SET is likely to be chosen more often than DET (H4).

METHODS

Description of the data

Data for this research were obtained from the RCT by Van Peperstraten et al. (2010) [32]. The choice for SET or DET should ideally be decided in a SDM process by an educated and empowered couple. In the RCT study of Van Peperstraten et al. (2010) a multifaceted strategy was used to promote SDM. To promote SDM, Van Peperstraten et al. developed a decision aid (DA) [see Box 2 for more information]. DAs are standardized, evidence-based tools intended to promote SDM [23]. Besides the evidence-based DA, this strategy consisted of support of an IVF nurse and reimbursement for an additional cycle of IVF for couples for whom the choice of SET caused a reduced chance of pregnancy [32]. In the Netherlands up to three IVF cycles are covered by the basic (but extensive) health insurance. The content of the DA and the reimbursement offer were discussed in person with a trained IVF nurse. All three elements of the strategy were provided before the counseling session that was part of standard care [32]. The control group received standard IVF care, including a session discussing the choice of SET or DET. Next to this standard care, the intervention group received the multifaceted empowerment strategy [32]. In the original RCT study, participating women completed three questionnaires (at inclusion, after intervention (but before starting treatment) and five weeks after embryo transfer) to measure decision-making outcomes and knowledge. Results showed that the proportion of couples in the intervention group who wanted to decide for themselves on the number of embryos transferred increased, while this percentage remained the same in the control group ($p<0.001$). Levels of both experienced knowledge ($p=0.001$) and actual knowledge ($p<0.001$) where higher in the intervention group compared to the control group [32]. For further detailed information see Van Peperstraten et al. (2010) [32].

Before the study, in 2005, 39% of the couples underwent SET after the first cycle [39]. The RCT was performed in five hospitals in the Netherlands. It included couples on the IVF waiting list between November 2006 and July 2007. The follow-up was continued until December 2008. Couples of women under 40 were included if they were on the waiting list for their first IVF cycle ever or a first cycle after a previous successful IVF. Couples were

excluded if SET was mandatory due to a strict medical indication. Written informed consent was provided by the couples before participation [32].

Selection of the data

In total, 308 couples at the beginning of their first IVF cycle were included in the intention to treat analysis (ITT) of Van Peperstraten et al. (2010) [32]. In all five hospitals approximately half of the couples received standard care, while the other half received the intervention. In this study, only couples that had the opportunity to choose between SET and DET were included. We, therefore, omitted from the 308 couples included in the ITT all couples: 1) where the woman was pregnant before starting IVF (N=20); 2) that never started IVF (N=13); and 3) that had none or just one embryo available (N=39 respectively N=14). Our sample included 222 couples, 113 in the control group and 109 in the intervention group, respectively. The outcome measure used in this study was the choice of either SET or DET. The data on this outcome were collected by Van Peperstraten et al. (2010) from local IVF registries [32]. Other variables included were whether a couple was involved in the intervention group or in the control group, and the hospital in which a couple was treated. In addition, we included four variables that are of medical relevance and might, therefore, affect the choice of SET or DET, and thus practice variation. For example, the older the woman is, the less likely she will become pregnant and the more likely she will have twins. The four variables included were: 1) the age of the woman (in years); 2) the duration of infertility (in years); 3) the presence of a good quality embryo (yes/no); and 4) any previous pregnancies (yes/no). Data on the presence of a good quality embryo were collected by Van Peperstraten et al. (2010) from local IVF registries [32]. Data for the other three variables were collected through a patient questionnaire which couples received when included to the study [32]. The woman's age and duration of infertility were calculated in this study on 31 December 2008 (= end of follow-up).

Statistical analyses of the data

We examined whether the control and intervention group were comparable for the characteristics included by performing descriptive statistics, and chi^2 tests (categorical variables) and t-tests (continuous variables) ($p < 0.05$). We then examined whether the five hospitals included did significantly differ with respect to the four variables that are of medical relevance. If there were differences between the five hospitals then we had to take these into account throughout the rest of our analyses, since these may have an impact on the choice of SET or DET. We performed descriptive statistics per hospital for the four variables. By chi^2 tests (categorical variables) and one-

way analyses of variance (ANOVA) (continuous variables), we tested if there were significant differences between the five hospitals for the woman's age, the duration of infertility, the presence of a good quality embryo, and for previous pregnancies ($p < 0.05$). If a significant difference was found between the hospitals for one of the aforementioned variables, we then performed an additional analysis to examine if there was an association between that variable and the outcome measure.

We then calculated, for each hospital, the percentage of couples that chose SET or DET, both in the control and in the intervention groups. We examined this in order to confirm that educated couples are inclined to choose SET. Next, we examined the variation **between** the hospitals. We first calculated for each hospital the percentage of SET in the control group and then in the intervention group. Thereafter, we calculated the range of SET percentages for the control groups and the intervention groups. A smaller range or difference between the highest and the lowest percentage of SET, implies less variation. Thus, if a strategy to promote SDM is associated with less variation between hospitals, then the range of SET percentages for the intervention group is smaller than that for the control group.

We now examined the variation **within** the hospitals by looking at the differences in variation between the control and the intervention group in each hospital. We calculated for each hospital the absolute difference between SET and DET in the control as well as in the intervention group. For example, if 40% chose SET and 60% DET, the absolute difference is 20. There is no variation if the proportion of SET to DET is 0% compared to 100%, or vice versa. Thus, there is no variation if the absolute difference is 100. On the other hand, the most variation is observed if the proportion of SET to DET is 50% to 50%. Thus, there is an absolute difference of 0. We can thus create a scale ranging from 0 to 100, where a score closer to 100 means less variation. We compared the scores of the control and the intervention group for each hospital. If the score in a hospital is higher in the intervention group than in the control group, and thus closer to 100 (= no variation), then a strategy to promote SDM is associated with less variation within that hospital. Complementary to the descriptive statistics, we performed a multilevel analysis (MLA) in MLwiN to examine variation between hospitals. A MLA takes into account the nested structure of the data as well as the differences in the number of patients per hospital. All statistical analyses were carried out using STATA, version 13.1.

RESULTS

Characteristics of the couples included

A description of the 222 couples included is given in Table 1. The number of couples included ranged from twelve couples in hospital five to 153 couples in hospital one (see Table 1). The control and intervention group did not differ significantly with respect to the characteristics included. Furthermore, no significant differences were observed between the five hospitals for the mean duration of infertility (p=0.256), the presence of a good quality embryo (p=0.406), and for previous pregnancies (p=0.403) (see Table 2). ANOVA showed a significant difference (p=0.032) for the variable, woman's age. An additional t-test showed no difference between woman's age and the choice of SET or DET (p=0.346). Thus, we decided not to include these four variables throughout the rest of the analyses.

Choice of SET

Table 3 shows the numbers and percentages of SET and DET for both the control and intervention groups, in total and per hospital. In total, 52% of the couples included in the intervention group chose SET, in comparison with 39% of the couples in the control group (p=0.046). To be more specific, in four of the five hospitals, couples in the intervention groups more often chose SET than DET. In hospital four, however, couples in the intervention group more often chose DET than SET (80% vs. 20%). Although in hospital four couples in the intervention group more often chose DET, they more often chose SET (20%) than couples in the control group (0%).

Variation between hospitals

The range of SET in both the control and intervention groups can also be observed in Table 3. The percentages of SET in the control groups ranged from 0.0% to 85.7%, while the percentages of SET in the intervention groups ranged from 20.0% to 87.5%. Therefore the range of SET is smaller in the intervention group than in the control group, which is an indication that a strategy to promote SDM reduced variation in the choice of SET or DET between hospitals. The MLA also indicated that the variation between hospitals was lower in the intervention group than in the control group. However, the difference was not significant.

Variation within hospitals

Figure 1 shows the differences in variation within hospitals by illustrating, per hospital, the absolute difference between SET and DET in the control group (standard care) and the intervention group (strategy to promote SDM). In one hospital (number 2) the absolute difference in the control group and the intervention group is the same. This means that the variation within hospital 2 is the same with or without a strategy to promote SDM. Within

hospital 3, the strategy to promote SDM appears to be associated with less variation, since the absolute difference in the control group is lower than in the intervention group (14 and 75, respectively). On the other hand, within the other three hospitals, numbers 1, 4 and 5, the strategy to promote SDM appears to be associated with more variation. Within these three hospitals the absolute difference in the control group is higher than in the intervention group (see Figure 1). Therefore, within some of the hospitals included, a strategy to promote SDM appears to be associated with more variation, while within other hospitals a strategy to promote SDM appears to be associated with less or the same level of variation.

DISCUSSION AND CONCLUSIONS

Principal findings

This study further elaborated upon and explored the association between SDM and variation in the choice of SET or DET both between and within hospitals. There was large variation in the choice of SET or DET between hospitals in the control group. Lower variation between hospitals was observed in the group with a strategy to promote SDM. Furthermore, we observed that within most hospitals the variation in the choice of SET or DET appeared to increase due to a strategy to promote SDM. This was particularly in hospitals where mainly DET was chosen in the control group.

What this study adds

Literature suggests that SDM reduces variation [23 29-31]. There was however, up to now, no clear evidence about this association. This study is the first that explored this association based on a case concerning the choice of SET or DET after IVF. We noticed that a strategy to promote SDM reduces variation between hospitals (confirming H1), while the variation within most hospitals appears to increase. The hypothesis in literature that SDM reduces variation is based on the observation that informed patients more often prefer less invasive treatments [3 26]. We found that in most hospitals couples in the intervention group more often chose SET. Although this does not imply that there will be less variation, since our results indicate that variation within most hospitals increased. This is because the level of variation without SDM differed between hospitals. For example, in some of the hospitals included mainly DET was preferred and there was thus almost no variation. Due to the strategy to promote SDM, however, SET was chosen more often, and thus the variation increased within such hospitals, since now both SET and DET are chosen (confirming H2). In one hospital SET was mainly chosen in the control group, and we therefore expected no change in variation, since we expected SET still preferred due to

the strategy to promote SDM (H3). However, we observed an increase in variation in that hospital, rejecting H3. In the two hospitals with the largest variation in the control group the variation decreased or remained equal, confirming H4. A subsequent implication is that an overall decrease in variation between hospitals, provides no indication about the change in variation within an individual hospital. Although based on a limited dataset, this study gives a first insight that SDM results in less variation between hospitals while suggesting another pattern of variation within hospitals.

Further research

This research focused on just one decision-making situation in obstetrics, and had only access to a limited dataset. The results, therefore, have to be interpreted with caution and further research is necessary both to underpin our results and to examine questions that remain unanswered. Nevertheless, our study provides a starting point for further empirical research within this area. For many medical problems no absolute best treatment option is available and so there are significant trade-offs among the available options [40 41]. We expect, however, that our results apply generally to medical problems with no absolute best treatment option. Decisions concerning such problems are defined as preference sensitive, since they depend on considerations of the benefits, disadvantages, and uncertainties of each treatment. For example some patients will prefer to accept a small risk of death in order to attempt to improve their function, while others will not [23 41]. Therefore, the best decisions cannot be made without including patients' preferences [40 41]. Well-known examples include chronic back pain, early-stage breast cancer and prostate cancer. For examples such as these it is believed that variation will change as a result of SDM. Future research has to confirm this by making use of data from multicentre RCT studies that applied intervention strategies (like a DA) to increase SDM in a specific consultation. Such RCT studies have been carried out [26], but have focused, comparable to the study we used, on outcome measures other than variation. Therefore, we decided in this study to perform a secondary analysis. Any possible multicentre studies should include a control and an intervention group which could thus measure actual treatment choices with and without SDM. This would allow researchers to examine whether SDM changes the pattern of variation, by, for example, using the same method as we did. Another possibility for further research is to conduct a new multicentre RCT study specifically aimed at analysing the relationship between SDM and medical practice variation.

Our results show that a strategy to promote SDM results in less variation between hospitals and indicates another pattern of variation within hospitals, confirming our hypotheses. These results appear to show that the decisions

made by informed patients have a pattern too. Choices made by informed patients appear to have a rate which deviates from baseline rates, irrespective of whether those are "low" or "high". This could be potentially named the informed patient rate. However, it can be questioned whether the rates we observed are indeed the informed patient rate, that is the results of what the couples want. It is possible that not all patients were able, or preferred, to take a shared decision about the choice of SET or DET. Nevertheless, results of the original RCT study show that levels of both experienced knowledge ($p=0.001$) and actual knowledge ($p<0.001$) where higher in the intervention group compared to the control group [32].Further research has to examine whether the actual choice was indeed the patients' preference and whether there are differences between groups of patients in this. In addition, a different pattern of variation due to SDM might be a positive indication for the quality of care. Good health care requires, among others, providing care that is respectful of, and responsive to, individual patients' preferences [19]. This is particularly true for preference sensitive decisions, since these decisions depend on patients' preferences regarding the benefits, disadvantages and uncertainties of each treatment. Further research has to examine whether SDM results in better quality of care for preference sensitive decisions.

The broad context of this study is about the influence of patients and their preferences on variation in medical practice. SDM is one option for including patients' preferences in medical decision-making. There are other options through which patients can express their preferences, and thus to influence the pattern of variation. For example, patients differ in how much pressure they are able to put on physicians [13 42]. They differ in their ability to take part in discussions over treatment with their physicians. Some patients are expected to be able to ask their physician for another treatment than, for example, the treatment that is recommended in a guideline or the one that is preferred by the physician [13 43]. If this is the case, then patients' preferences appear to influence the treatment chosen, and thus the variation. Further research is recommended into these situations.

The strengths and weaknesses of the study

A strength of our study is that we are the first to elaborate further upon and explore the association between SDM and variation in medical treatment. We examined this association to get insight into whether including patients' preferences through a strategy to promote SDM results in another pattern of variation in medical treatment. Another strength is the use of data from a RCT. We had the opportunity to compare the variation in the choice of SET or DET with, and without, SDM. It might be possible that the choice of SET or DET in the control group is influenced by physicians, since they treated both couples in the intervention and in the control group. Ideally, data

would have been available about the percentage of SET and DET before the RCT, allowing us to compare the intervention and control group with these percentages. Though another study showed that in 2005, before the RCT, 39% of twin prone couples in two Dutch hospitals chose SET [39], which is comparable to the percentage of SET in the control group. It seems plausible to use the control group as the situation before SDM. We performed descriptive statistics to analyse our data, because of the low numbers of couples included in the hospitals. We have taken into account the nested structure of the data by performing our analyses per hospital. However, we did not take into account the differences in the number of patients per hospital. We therefore also performed an MLA to examine the variation between hospitals. The MLA supported the results of the descriptive statistics, however, the difference was not significant. From the dataset it was also known only in which hospital a couple was treated, but not by which physician within that hospital. However, only one or two physicians per hospital treated all couples in that hospital so we do not expect that this will affect our conclusions. Further research should ideally be performed with a larger dataset, preferably using multilevel analysis in order to test the hypothesis of this study. This would acknowledge that patients are nested hierarchically within physicians and physicians within hospitals. In addition, further research has to include in the analyses socio-demographics that might have an influence on the treatment decision. Another limitation might be that the original RCT did not assessed whether more SDM actually took place. As such, it is unclear whether the strategy really led to more SDM. A final limitation might be that the intervention consisted of different elements, and thus it is difficult to assess separately the effects of these elements. Despite the fact that at the end of the follow-up period only 4% of the couples qualified for reimbursement of a fourth cycle, reimbursement might have played a role in the decision [32]. However, a follow-up study showed that – compared to the other elements – the reimbursement offer was rated least important by the couples in choice for SET or DET [44].

Conclusions

This study was the first to elaborate further upon and explore the relationship between including patients' preferences in medical decision-making and practice variation. Although based on a limited dataset, our study gives a first insight that including patients' preferences through SDM results in less variation between hospitals, and indicates another pattern of variation within hospitals. The variation that results from patient preferences could be potentially named the informed patient rate. The results of this study provide the starting point for further empirical research within this area.

Acknowledgements We thank Peter Spreeuwenberg (NIVEL) for advising on and performing the MLA.

Authorship AB was involved in the design of the study, performed the statistical analyses, and drafted the manuscript. JDJ and LVD were involved in the design of the study, assisted in interpreting the results, assisted in drafting the manuscript, and revised the manuscript. PG was involved in the design of the study, and revised the manuscript. AVP was the primary author of the original RCT and was therefore responsible for the data collection, and revised the manuscript. All authors have read and approved the final manuscript.

Conflict of interest All authors have completed the unified competing interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare (1) all no financial support for the work submitted from anyone other than their employer; (2) AB, PG, AVP and JDJ no financial relationships with commercial entities that might have an interest in the work submitted. LVD received unrestricted grants from Bristol-Myers Squibb, Pfizer and Astra Zeneca for studies not related to this study; (3) all no spouses, partners, or children with relationships with commercial entities that might have an interest in the work submitted; and (4) all no non-financial interests that may be relevant to the submitted work.

Source of funding The RCT study was funded by the Netherlands Organisation for Health Research and Development (grant no 945-16-105). For this secondary analysis of the RCT no additional funding was obtained.

Ethics approval The RCT study was approved by the regional ethics committee for medical research.

Data sharing statement The dataset of the RCT study is available upon request from AVP.

Transparency All authors are a guarantor and affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no imported aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

REFERENCES

1. Paul-Shaheen P, Clark JD, Williams D. Small area analysis: a review and analysis of the North American literature. J Health Polit Policy Law 1987;**12**(4):741-809.

2. Wennberg JE. Understanding geographic variations in health care delivery. NEJM 1999;**340**(1):52-3.

3. Wennberg JE. *Tracking medicine: a researcher's quest to understand health care*: Oxford University Press, 2010.

4. Corallo AN, Croxford R, Goodman DC, et al. A systematic review of medical practice variation in OECD countries. Health Policy 2014;**114**(1):5-14.

5. Emanuel EJ, Emanuel LL. Four models of the physician-patient relationship. JAMA 1992;**267**(16):2221-6.

6. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med 1997;**44**(5):681-92.

7. Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. BMJ 2002;**325**(7370):961.

8. McPherson K. Why do variations occur? In: Andersen TF, Mooney G, eds. The challenges of medical practice variations. London: The Macmillan Press, 1990:16-35.

9. Wennberg JE, Gittelsohn A. Health care delivery in Maine I: patterns of use of common surgical procedures. J Maine Med Assoc 1975;**66**(5):123-30, 49.

10. Mercuri M, Gafni A. Medical practice variations: what the literature tells us (or does not) about what are warranted and unwarranted variations. J Eval Clin Pract 2011;**17**(4):671-7.

11. De Jong JD, Groenewegen PP, Westert GP. Sociological Model for Understanding Medical Practice Variations. In: Sobolev B (Series ed) JA, Stukel T (eds). , ed. Health Services Research Series Medical Practice Variations. New York: Springer, 2015.

12. Chassin MR. Explaining geographic variations. The enthusiasm hypothesis. Med Care 1993;**31**(5 Suppl):Ys37-44.

13. De Jong JD. Explaining medical practice variation. Social organization and institutional mechanisms. NIVEL, 2008.

14. Westert GP, Groenewegen PP. Medical practice variations: changing the theoretical approach. Scand J Public Health 1999;**27**(3):173-80.

15. van de Bovenkamp HM, Trappenburg MJ, Grit KJ. Patient participation in collective healthcare decision making: the Dutch model. Health Expect 2010;**13**(1):73-85.

16. Guadagnoli E, Ward P. Patient participation in decision-making. Soc Sci Med 1998;**47**(3):329-39.

17. Frosch DL, Kaplan RM. Shared decision making in clinical medicine: past research and future directions. Am J Prev Med 1999;**17**(4):285-94.

18. Fraenkel L. Incorporating patients' preferences into medical decision making. Med Care Res Rev 2013;**70**(1 Suppl):80s-93s.

19. Institute of Medicine. *Crossing the quality chasm: A new health system for the 21st century*: National Academies Press, 2001.

20. Wennberg JE. *Variation in Use of Medicare Services among Regions and Selected Academic Medical Centers: Is More Better?*: Commonwealth Fund, 2005.

21. Greer AL, Goodwin JS, Freeman JL, et al. Bringing the patient back in. Guidelines, practice variations, and the social context of medical practice. Int J Technol Assess Health Care 2002;**18**(4):747-61.

22. Charles C, Gafni A, Whelan T. Decision-making in the physician-patient encounter: revisiting the shared treatment decision-making model. Soc Sci Med 1999;**49**(5):651-61.

23. O'Connor AM, Llewellyn-Thomas HA, Flood AB. Modifying unwarranted variations in health care: shared decision making using patient decision aids. Health Affairs 2004:Var63-72.

24. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012;**27**(10):1361-7.

25. Elwyn G, Laitner S, Coulter A, et al. Implementing shared decision making in the NHS. BMJ 2010;**341**:c5146.

26. Stacey D, Bennett CL, Barry MJ, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2011(10):Cd001431.

27. Deyo RA, Cherkin DC, Weinstein J, et al. Involving patients in clinical decisions: impact of an interactive video program on use of back surgery. Med Care 2000;**38**(9):959-69.

28. McClure W. Toward development and application of a qualitative theory of hospital utilization. Inquiry 1982;**19**(2):117-35.

29. O'Connor AM, Wennberg JE, Legare F, et al. Toward the 'tipping point': decision aids and informed patient choice. Health Affairs 2007;**26**(3):716-25.

30. Stiggelbout AM, Van der Weijden T, De Wit MP, et al. Shared decision making: really putting patients at the centre of healthcare. BMJ 2012;**344**:e256.

31. Weinstein JN, Bronner KK, Morgan TS, et al. Trends and geographic variations in major surgery for degenerative diseases of the hip, knee, and spine. Health Affairs 2004:Var81-9.

32. van Peperstraten A, Nelen W, Grol R, et al. The effect of a multifaceted empowerment strategy on decision making about the number of embryos transferred in in vitro fertilisation: randomised controlled trial. BMJ 2010;**341**:c2501.

33. Pandian Z, Marjoribanks J, Ozturk O, et al. Number of embryos for transfer following in vitro fertilisation or intra-cytoplasmic sperm injection. Cochrane Database Syst Rev 2013;**7**:Cd003416.

34. Pinborg A. IVF/ICSI twin pregnancies: risks and prevention. Hum Reprod Update 2005;**11**(6):575-93.

35. Andersen AN, Goossens V, Ferraretti AP, et al. Assisted reproductive technology in Europe, 2004: results generated from European registers by ESHRE. Hum Reprod 2008;**23**(4):756-71.

36. de Mouzon J, Goossens V, Bhattacharya S, et al. Assisted reproductive technology in Europe, 2006: results generated from European registers by ESHRE. Hum Reprod 2010;**25**(8):1851-62.

37. Ferraretti AP, Goossens V, Kupka M, et al. Assisted reproductive technology in Europe, 2009: results generated from European registers by ESHRE. Hum Reprod 2013;**28**(9):2318-31.

38. van Peperstraten A. Implementation of single embryo transfer: a patient directed strategy. Radboud Universiteit Nijmegen, 2009.

39. Van Peperstraten AM, Kreuwel IA, Hermens RP, et al. Determinants of the choice for single or double embryo transfer in twin prone couples. Acta Obstet Gynecol Scand 2008;**87**(2):226-31.

40. Barry MJ. Shared decision making: informing and involving patients to do the right thing in health care. J Ambul Care Manage 2012;**35**(2):90-8.

41. Dartmouth Atlas. A Dartmouth Atlas Project Topic Brief Preference-Sensitive Care. Lebanon: The Dartmouth Atlas of Health Care, 2007.

42. Westert GP, Groenewegen PP, Lindenberg SM. De invloed van ziekenhuispatiënten op de duur van de hospitalisatie. TSG 1991;**69**:184-94.

43. Hajjaj FM, Salek MS, Basra MK, et al. Non-clinical influences on clinical decision-making: a major challenge to evidence-based practice. J R Soc Med 2010;**103**(5):178-87.

44. Kreuwel IA, van Peperstraten AM, Hulscher ME, et al. Evaluation of an effective multifaceted implementation strategy for elective single-embryo transfer after in vitro fertilization. Hum Reprod 2013;**28**(2):336-42.

45. No authors listed. Prevention of twin pregnancies after IVF/ICSI by single embryo transfer. ESHRE Campus Course Report. Hum Reprod 2001;**16**(4):790-800.

46. Wolner-Hanssen P, Rydhstroem H. Cost-effectiveness analysis of in-vitro fertilization: estimated costs per successful pregnancy after transfer of one or two embryos. Hum Reprod 1998;**13**(1):88-94.

47. Lukassen HG, Schonbeck Y, Adang EM, et al. Cost analysis of singleton versus twin pregnancies after in vitro fertilization. Fertil Steril 2004;**81**(5):1240-6.

48. van Peperstraten AM, Hermens RP, Nelen WL, et al. Deciding how many embryos to transfer after in vitro fertilisation: development and pilot test of a decision aid. Patient Educ Couns 2010;**78**(1):124-9.

49. Elwyn G, O'Connor A, Stacey D, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ 2006;**333**(7565):417.

BOXES

Box 1: The choice of single or double embryo transfer: A complex decision-making problem

The choice of a single or double embryo transfer after IVF is a complex decision-making problem because of the need to find a balance between the risk of complications of multiple birth and the best chance of pregnancy [32]. Some sub-fertile couples and professionals regard twin pregnancies as a success, however, they could also be considered as a side effect or even a complication [45]. Twin pregnancies are associated with higher morbidity and mortality rates for both mother and child compared to singleton pregnancies [34]. Moreover, complications of twin pregnancies cause substantial use of medical budgets [46 47]. Subsequently, twin pregnancies are increasingly regarded as undesirable. To prevent twin pregnancies professionals and couples could choose single embryo transfer (SET) instead of double embryo transfer (DET) [34 45]. However, this may be disadvantageous, since it could result in a lower pregnancy rate per IVF cycle [33]. The choice of SET or DET is ideally decided through SDM [48].

Box 2: The choice of single or double embryo transfer: An evidence-based decision aid

Van Peperstraten et al. (2010) developed and tested the evidence-based decision aid (DA) for deciding how many embryos to transfer during IVF [48]. The DA was developed according to the checklist of the International Patient Decision Aids Standards Collaboration, which consists of 50 items divided between three domains, content, development, and effectiveness [48 49]. The purpose of the DA is to give couples all the information needed to make the choice to transfer one or two embryos and to relate the information to their own personal situation. The DA consists of three chapters: 1) information about the chances of a single pregnancy or a twin pregnancy; 2) information about the risks of twin pregnancies; and 3) an explanation of the available options and an action plan [48]. The DA is available in English at: www.umcn.nl/ivfda-en.

TABLES

Table 1. Characteristics of the couples included (N = 222). Values are numbers unless otherwise stated.

Characteristics	Control group (N = 113)	Intervention group (N = 109)	Total (N = 222)	p-value[c]
Hospital				
hospital 1	79	74	153	N.A.
hospital 2	7	7	14	
hospital 3	7	8	15	
hospital 4	13	15	28	
hospital 5	7	5	12	
Mean (sd) age of woman (years)[a]	33.9 (3.85) (range 21-41 years)	33.5 (3.88) (range 25-41 years)	33.7 (3.86) (range 21-41 years)	0.475
Mean (sd) duration of infertility (years)[a]	4.03 (2.08) (range 1-13 years) (N=101)	3.94 (1.91) (range 1-12 years) (N=98)	3.98 (2.00) (range 1-13 years) (N=199)	0.749
Presence of a good quality embryo				
no	41	28	69	0.088
yes	72	81	153	
Previous pregnancies[b]	(N=113)	(N=108)	(N=221)	
no	63	63	126	0.698
yes	50	45	95	

[a] Calculated on December 31, 2008 based on information filled out in patients' questionnaires. As a result, we have a higher mean for age and duration of infertility than Van Peperstraten et al., BMJ, 2010.

[b] Based on the question: "Have you ever been pregnant?"

[c] $p < 0.05$ is significant

Table 2. Characteristics of the couples included per hospital.

Characteristics	Total (N = 222)	p-value[c]
Mean (sd) age of woman (years)[a]	33.7 (3.86) (21-41)	
hospital 1 (N = 153)	33.8 (3.63) (21-41)	**0.032**
hospital 2 (N = 14)	30.9 (4.54) (25-39)	
hospital 3 (N = 15)	34.6 (3.96) (28-40)	
hospital 4 (N = 28)	33.4 (4.53) (25-41)	
hospital 5 (N = 12)	35.2 (2.86) (30-38)	
Mean (sd) duration of infertility (years)[a]	4.0 (2.00) (1-13)	
hospital 1 (N = 139)	4.1 (2.20) (1-13)	0.256
hospital 2 (N = 11)	4.0 (1.26) (2-6)	
hospital 3 (N = 11)	4.3 (1.85) (2-8)	
hospital 4 (N = 26)	3.2 (1.22) (1-6)	
hospital 5 (N = 12)	4.0 (1.13) (2-6)	
Presence of a good quality embryo (% yes)	68.9%	
hospital 1 (N = 153)	72.6%	0.406
hospital 2 (N = 14)	64.3%	
hospital 3 (N = 15)	66.7%	
hospital 4 (N = 28)	60.7%	
hospital 5 (N = 12)	50.0%	
Previous pregnancies[b] (% yes)	43.0%	
hospital 1 (N = 153)	45.8%	0.403
hospital 2 (N = 14)	21.4%	
hospital 3 (N = 14)	50.0%	
hospital 4 (N = 28)	39.3%	
hospital 5 (N = 12)	33.3%	

[a] Calculated on December 31, 2008 based on information filled out in patients' questionnaires. As a result, we have a higher mean for age and duration of infertility than Van Peperstraten et al., BMJ, 2010.

[b] Based on the question: "Have you ever been pregnant?"

[c] $p < 0.05$ is significant

Table 3. The choice of SET or DET total group, and per hospital. Values are numbers (%) unless otherwise stated.

	Control group	Intervention group	p-value[a]
Total			
SET	44 (38.9%)	57 (52.3%)	**0.046**
DET	69 (61.1%)	52 (47.7%)	
Hospital 1			
SET	31 (39.2%)	40 (54.1%)	0.066
DET	48 (60.8%)	34 (46.0%)	
Hospital 2			
SET	3 (42.9%)	4 (57.1%)	0.593
DET	4 (57.1%)	3 (42.9%)	
Hospital 3			
SET	4 (57.1%)	7 (87.5%)	0.185
DET	3 (42.9%)	1 (12.5%)	
Hospital 4			
SET	0 (0.0%)	3 (20.0%)	0.088
DET	13 (100.0%)	12 (80.0%)	
Hospital 5			
SET	6 (85.7%)	3 (60.0%)	0.310
DET	1 (14.3%)	2 (40.0%)	

[a] $p < 0.05$ is significant

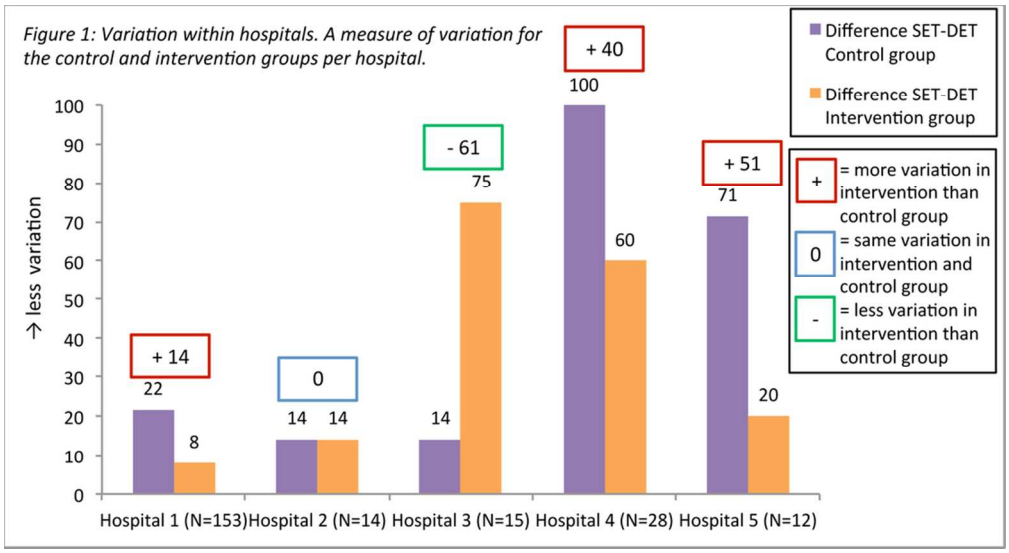


Figure 1: Variation within hospitals. A measure of variation for the control and intervention groups per hospital.
94x51mm (300 x 300 DPI)

CONSORT 2010 checklist of information to include when reporting a randomised trial*

We performed a secondary analysis of a RCT. The original RCT was published in the BMJ (2010) as "The effect of a multifaceted empowerment strategy on decision making about the number of embryos transferred in in vitro fertilisation: randomised controlled trial" by Van Peperstraten et al. We included this article as supplement to our submission.

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	See Van Peperstraten et al. (2010)
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	Page 2 of this submission
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	Page 4-5 of this submission
	2b	Specific objectives or hypotheses	Page 4-5 of this submission
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	See Van Peperstraten et al. (2010)
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	NA
Participants	4a	Eligibility criteria for participants	See Van Peperstraten et al. (2010) and Page 6 of this submission
	4b	Settings and locations where the data were collected	See Van Peperstraten et al. (2010) and Page 6 of this submission
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	See Van Peperstraten et al. (2010) and Page 6 of this submission
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	Page 6-7 of this submission

	6b	Any changes to trial outcomes after the trial commenced, with reasons	NA
Sample size	7a	How sample size was determined	See Van Peperstraten et al. (2010) en Page 6 of this submission
	7b	When applicable, explanation of any interim analyses and stopping guidelines	NA
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	See Van Peperstraten et al. (2010)
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	See Van Peperstraten et al. (2010)
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	See Van Peperstraten et al. (2010)
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	See Van Peperstraten et al. (2010)
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	NA
	11b	If relevant, description of the similarity of interventions	NA
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	Page 7-8 of this submission
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	NA
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	Table 1 and 2 and 3 of this submission
	13b	For each group, losses and exclusions after randomisation, together with reasons	Page 6 of this submission
Recruitment	14a	Dates defining the periods of recruitment and follow-up	See Van Peperstraten et al. (2010) and Page 6 of this submission
	14b	Why the trial ended or was stopped	NA
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Table 1 and 2 of this submission
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Table 1 and 2 and 3 of this submission
Outcomes and	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its	Table 3 and Figure 1 and

estimation		precision (such as 95% confidence interval)	Page 8-9 of this submission
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	Table 3 and Figure 1
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	NA
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	NA
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	See Van Peperstraten et al. (2010) and Page 12 of this submission
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Page 10 of this submission
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	Page 10-12 of this submission
Other information			
Registration	23	Registration number and name of trial registry	See Van Peperstraten et al. (2010) and Page 2 of this submission
Protocol	24	Where the full trial protocol can be accessed, if available	NA
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	See Van Peperstraten et al. (2010) and Page 13 of this submission

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.