

Hysterectomy, endometrial destruction, and levonorgestrel releasing intrauterine system (Mirena) for heavy menstrual bleeding: systematic review and meta-analysis of data from individual patients

L J Middleton,¹ R Champaneria,¹ J P Daniels,¹ S Bhattacharya,² K G Cooper,³ N H Hilken,¹ P O'Donovan,^{4,5} M Gannon,⁶ R Gray,¹ K S Khan,⁷ on behalf of the International Heavy Menstrual Bleeding Individual Patient Data Meta-analysis Collaborative Group

EDITORIAL by Clarke

¹Birmingham Clinical Trials Unit, University of Birmingham, Birmingham B15 2TT

²Section of Applied Clinical Sciences, Division of Applied Health Sciences, School of Medicine and Dentistry, University of Aberdeen, Aberdeen Maternity Hospital, Foresterhill, Aberdeen AB25 2ZD

³Department of Obstetrics and Gynaecology, Aberdeen Royal Infirmary, Foresterhill, Aberdeen AB25 2ZN

⁴Bradford Royal Infirmary (Bradford Teaching Hospitals NHS Foundation Trust), Bradford BD6 6RJ

⁵University of Bradford, Bradford BD7 1DP

⁶Department of Obstetrics and Gynaecology, Midland Regional Hospital, Mullingar, Co Westmeath, Republic of Ireland

⁷Academic Department of Obstetrics and Gynaecology, Division of Reproductive and Child Health, University of Birmingham, Birmingham Women's Hospital, Birmingham B15 2TG

Correspondence to: L Middleton
l.j.middleton@bham.ac.uk

Cite this as: *BMJ* 2010;341:c3929
doi: 10.1136/bmj.c3929

This is a summary of a paper that was published on *bmj.com* as *BMJ* 2010;341:c3929

STUDY QUESTION How dissatisfied are women after treatment for heavy menstrual bleeding with hysterectomy, first and second generation endometrial destruction techniques, or the levonorgestrel releasing intrauterine system (Mirena)?

SUMMARY ANSWER First and second generation endometrial destruction techniques were associated with greater dissatisfaction than hysterectomy, although rates of dissatisfaction were low after all treatments.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS Less invasive alternatives to hysterectomy for the treatment of heavy menstrual bleeding, such as endometrial destruction and Mirena, have become increasingly popular. More women were dissatisfied with treatment after first or second generation endometrial destruction techniques than after hysterectomy, but hysterectomy is associated with increased length of hospital stay and recovery period.

Selection criteria for studies

We sought data on individual patients from randomised controlled trials comparing hysterectomy, first and second generation endometrial destruction techniques, and Mirena in women with heavy menstrual bleeding in whom other medical treatments had failed. Published aggregate data were included if individual data could not be attained. Studies were identified with a comprehensive search of the Cochrane Library, Medline, Embase, and CINAHL databases, reference lists, and contact with experts. Indirect comparisons were made on the primary outcome measure if a direct comparison was not available.

Primary outcome

Dissatisfaction with treatment at around 12 months.

Main results and role of chance

Raw data were available from 2814 women randomised into 17 trials (seven trials including 1359 women for first versus second generation endometrial destruction; six trials including 1042 women for hysterectomy versus first generation endometrial destruction; one trial including 236 women for hysterectomy versus Mirena; three trials including 177 women for second generation endometrial destruction versus Mirena).

At around 12 months, more women were dissatisfied with outcome after first generation hysteroscopic techniques than after hysterectomy (13% v 5%, odds ratio 2.5, 95% confidence interval 1.5 to 3.9, $P < 0.001$), but hospital stay (weighted mean difference 3.0 days, 2.9 to 3.1 days, $P < 0.001$) and time to resumption of normal activities (5.2 days, 4.7 to 5.7 days, $P < 0.001$) were longer for hysterectomy. Unsatisfactory outcomes were comparable with first and second generation techniques (odds ratio 1.2, 0.88 to 1.6, $P = 0.2$), though second generation techniques were quicker (weighted mean difference 14.5 minutes, 13.7 to 15.3 minutes, $P < 0.001$) and women recovered sooner (0.48 days, 0.20 to 0.75 days, $P < 0.001$) with fewer procedural complications. Indirect comparison suggested more unsatisfactory outcomes with second generation techniques than with hysterectomy (11% v 5%; odds ratio 2.3, 1.3 to 4.2, $P = 0.006$). Similar estimates were seen when Mirena was indirectly compared with hysterectomy (17% v 5%; odds ratio 2.2, 0.94 to 5.3, $P = 0.07$), although this comparison lacked power because of the limited amount of data available for analysis.

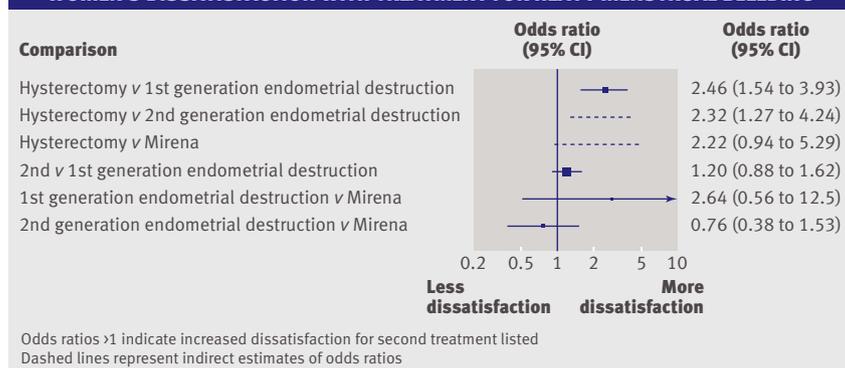
Bias, confounding, and other reasons for caution

The review's inferences were limited by inconsistent measures across the trials; studies involving endometrial destruction and Mirena focused on comparing reduction in bleeding, while hysterectomy trials focused on quality of life and use of resources. Although we have shown that satisfaction rates are closely related to quality of life, a disease specific quality of life tool would have been a preferable primary measure for meta-analysis. This outcome was not assessed in most of the studies identified. The current body of evidence comparing Mirena was limited to several small studies with high levels of non-compliance and prohibited us from making any strong conclusions about this treatment.

Study funding/potential competing interests

This review was funded by the Health Technology Assessment Programme of the National Institute for Health Research (05/45/02).

WOMEN'S DISSATISFACTION WITH TREATMENT FOR HEAVY MENSTRUAL BLEEDING



Inconsistent reporting of surrogate outcomes in randomised clinical trials: cohort study

Jeppe Lerche la Cour,¹³ Jesper Brok,¹² Peter C Gøtzsche³

¹Copenhagen Trial Unit, Rigshospitalet, Copenhagen University Hospital, DK-2100 Copenhagen, Denmark

²Pediatric Department, Hvidovre Hospital, Copenhagen

³Nordic Cochrane Centre, Rigshospitalet and University of Copenhagen

Correspondence to: J L la Cour Jeppelerche@hotmail.com

Cite this as: *BMJ* 2010;**341**:c3653
doi: 10.1136/bmj.c3653

This is a summary of a paper that was published on bmj.com as *BMJ* 2010;**341**:c3653

STUDY QUESTION

Do authors of randomised clinical trials convey the fact that they have used a surrogate outcome and do they discuss the surrogate's validity?

SUMMARY ANSWER

About one in five published randomised clinical trials uses a surrogate as a primary outcome, and only about one third of these trials report adequately on the use of a surrogate and its validity.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS

Uncritical use of surrogate outcomes can be misleading and may have resulted in implementation of harmful interventions. Better reporting on surrogate outcomes is needed to avoid misleading conclusions and uncritical acceptance of new treatments.

Design

We retrieved all randomised clinical trials published in 2005 and 2006 in six major general medical journals, which had used a surrogate as a primary outcome. An outcome was classified as a surrogate if it did not directly measure how a patient feels, functions, or survives (for example, bone mineral content). JLC screened all abstracts for eligibility, and JB or PCG reassessed eligibility. Two authors independently extracted data on whether the author reported use of a surrogate and whether the validity of the surrogate was discussed.

Main results

We included 109 randomised clinical trials with a primary surrogate outcome (17% of all such trials from the included journals). In 62 of the trials (57%, 95% confidence interval 47% to 67%) the authors clearly conveyed that they had used a surrogate primary outcome. Only 38 trials (35%, 26% to

45%) also discussed the surrogate's validity. Neither trials sponsored by for profit companies nor trials recommending the experimental treatment were significantly more likely to report inadequately on surrogates (relative risk 0.67, 95% confidence interval 0.40 to 1.10 and 1.93, 0.90 to 4.14).

Bias, confounding, and other reasons for caution

Only one author screened all trial abstracts for eligibility. This may have led to a few relevant trials using borderline surrogates being overlooked, but clearcut cases were not likely to have been missed. Two authors independently assessed the trials for inclusion. In the case of no consensus, the trial was not included. This raises the possibility that the prevalence of trials using surrogate outcomes is marginally higher than we found. It is not always clear whether trialists report adequately on the use and validity of surrogates. Furthermore, outcomes such as length of hospital stay, weight loss, and incidence of cancer can be both clinical and surrogate outcomes depending on the trial set-up. However, two authors independently extracted the data and only when consensus could be reached was the trial included and data used, minimising the degree of subjectivity.

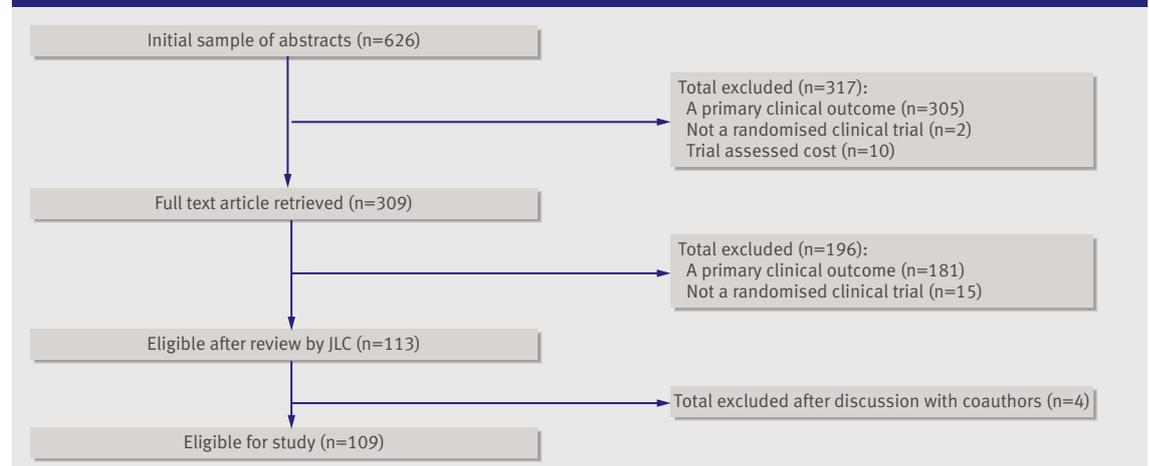
Generalisability

We chose six major journals because they have a high impact on clinical decisions and because they do not focus on a single discipline like specialist journals do. Hence readers cannot be expected to have prior knowledge on surrogates beyond their own clinical area of expertise. This strengthens the need to mention when a surrogate has been used and to discuss its validity.

Study funding

All authors worked independently of the funders. We have no competing interests.

FLOW OF PAPERS THROUGH STUDY



Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review

Gloria Cordoba,¹ Lisa Schwartz,² Steven Woloshin,² Harold Bae,² Peter C Gøtzsche¹

EDITORIAL by Freemantle and Calvert

¹Nordic Cochrane Centre, Rigshospitalet and University of Copenhagen, Blegdamsvej 9, DK-2100 Copenhagen, Denmark

²Dartmouth Institute for Health Policy, Dartmouth Medical School, New Hampshire, USA

Correspondence to: P C Gøtzsche pcg@cochrane.dk

Cite this as: *BMJ* 2010;341:c3920
doi: 10.1136/bmj.c3920

This is a summary of a paper that was published on bmj.com as *BMJ* 2010;341:c3920

STUDY QUESTION

Are composite outcomes, which have combined several components into a single measure, appropriately defined, reported, and interpreted?

SUMMARY ANSWER No. The use of composite outcomes in clinical trials is highly problematic, with components often unreasonably combined, inconsistently defined, and inadequately reported.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS

Composite outcomes in clinical trials are theoretically attractive, as they reduce sample size requirement, costs, and time. However, they should generally be avoided, as their use leads to much confusion and bias.

Selection criteria for studies

A PubMed search identified parallel group, randomised clinical trials published in 2008 that had a primary composite outcome.

Primary outcomes

Definition, reporting, and interpretation of composite outcomes.

Main results and role of chance

Of 40 included trials, 29 (73%) were about cardiovascular

topics, and 24 (60%) were entirely or partly industry funded. The composite outcomes had a median of three components (range 2–9). Death or cardiovascular death was the most important component in 33 trials (83%). Only one trial provided a good rationale for the choice of components. We judged that the components were not of similar importance in 28 trials (70%)—in 20 of these, death was combined with hospital admission. Other major problems were changes in the definition of the composite outcome between the abstract, methods, and results sections (13 trials); missing, ambiguous, or uninterpretable data (nine trials); and post hoc construction of composite outcomes (four trials). Only 24 trials (60%) provided reliable estimates for both the composite and its components, and only six trials (15%) had components of similar, or possibly similar, clinical importance, and provided reliable estimates. In 11 of 16 trials with a statistically significant result for the composite outcome, the abstract conclusion falsely implied that the effect also applied to the most important component.

Bias, confounding, and other reasons for caution

Our results may not be generalisable outside the cardiovascular area.

Study funding/potential competing interests

No specific funding for this study, and no competing interests declared.

DEFINITION AND REPORTING OF COMPOSITE OUTCOMES IN 40 TRIALS

	No (%) of trials
Composite definition	
Most important component:	
Death (all causes or disease specific)	33 (83)
Clinical outcome (hospital admission or symptom)	7 (17)
Clinical importance of components:	
Similar	7 (18)
Might be similar	5 (13)
Not similar	28 (70)
Author discussion of composite outcome:	
No discussion	33 (83)
Explains rationale for composite	1 (3)
Acknowledges problems with composite	6 (15)
Reporting of composite	
Components consistent between abstract, methods, and results	27 (68)
Components inconsistent:	13 (33)
Major inconsistency (components added or deleted)	5 (13)
Minor inconsistency (ambiguous wording change)	8 (20)
Data for components provided	31 (78)
Data for components not provided:	9 (23)
Missing data	2 (5)
Ambiguous data	7 (18)

Predicting which people with psychosocial distress are at risk of becoming dependent on state benefits: analysis of routinely available data

Will Whittaker,¹ Matt Sutton,¹ Margaret Maxwell,² Rosalia Munoz-Arroyo,³ Sara MacDonald,⁴ Andrew Power,⁵ Michael Smith,⁶ Philip Wilson,⁴ Jill Morrison⁴

EDITORIAL by Gabbay

¹Health Sciences Research Group, School of Community Based Medicine, University of Manchester, Jean McFarlane Building, Manchester M13 9PL

²Nursing, Midwifery and AHP Research Unit, University of Stirling, Stirling FK9 4LA

³Information Services Division, NHS National Services Scotland, Area 122A, Edinburgh EH12 9EB

⁴Division of Community Based Sciences, Faculty of Medicine, University of Glasgow, Glasgow G12 9LX

⁵Pharmacy and Prescribing Support Unit, Queens Park House, Victoria Infirmary, Glasgow G42 9TT

⁶Department of Psychological Medicine, Division of Community Based Sciences, Gartnavel Royal Hospital, Glasgow G12 0XH

Correspondence to: J Morrison jmm4y@clinmed.gla.ac.uk

Cite this as: *BMJ* 2010;**341**:c3838
doi: 10.1136/bmj.c3838

This is a summary of a paper that was published on bmj.com as *BMJ* 2010;**341**:c3838

STUDY QUESTION

Can we predict which people with psychological distress are at increased risk of becoming dependent on long term sickness related state benefits?

SUMMARY ANSWER

There were significantly increased rates of caseness according to the general health questionnaire for two years before, and frequent consultations for three years before, patients went on to receive incapacity benefit.

WHAT IS KNOWN AND WHAT THIS STUDY ADDS

Mild mental disorders and frequent consultations in general practice are known to be associated with long term sickness absence. These factors can also predict long term sickness absence up to three years before people become dependent on benefit.

Participants and setting

Participants in the British Household Panel Surveys, 1991-2007 (1738 incapacity benefit claimants), and the Scottish Health Surveys in 1995, 1998, and 2003 (7932, 12939, and 11472 respondents, respectively).

Design

Data were collected for a maximum of 16 years before the start of the first claim for incapacity benefit and a maximum of 15 years after the start of the first claim. To test for significant changes in rates of caseness according to the general health questionnaire (GHQ-12) and frequent consultations with a general practitioner during receipt of incapacity benefit, we used confidence intervals from random effects logistic regression on year dummies with clustering by respondent. To test for differences in rates of claims across general practices we modelled a logistic regression for work incapacity, with clustering by general practice.

Main results and role of chance

There was a significant increase in rates of caseness from two years before starting to claim incapacity benefit (odds ratio 1.6, 95% confidence interval 1.3 to 1.9) and an increase in frequent consultation from three years before starting to claim (1.8, 1.3 to 2.4). People with GHQ-12 caseness showed a significant increase in frequent consultations from two years before they started to claim incapacity benefit (2.1, 1.4 to 3.2). There was a small (intracluster correlation coefficient 0.010, 0.001 to 0.061) and insignificant ($P=0.13$) amount of variation across general practices in Scotland in rates of claiming incapacity benefit after adjustment for other explanatory variables.

Bias, confounding, and other reasons for caution

We used a cut off of 10 or more consultations to describe frequent consulters as this was the highest level used in the British Household Survey, but recent research has suggested the use of a top decile cut off for defining frequent attendance might have more discriminative power. The GHQ-12 gives an indication of potential mental health problems, rather than providing a definite diagnosis. We modelled only one episode of being in receipt of incapacity benefit per individual, and this had to have started during the sample period. We excluded anyone already in receipt of incapacity benefit in 1990-1. We did not know the main cause of transition on to incapacity benefit, so it might have been because of another condition, such as a physical illness.

Generalisability to other populations

Similar results would probably be found in other countries with comparable systems of providing social benefits for people who are unable to work because of ill health.

Study funding/potential competing interests

This study was funded by a project grant from the Chief Scientist Office of the Scottish Government Health Directorate.

RATES OF FREQUENT CONSULTATIONS WITH GENERAL PRACTITIONER IN PEOPLE WITH MENTAL HEALTH PROBLEMS WHO HAVE A PERIOD ON INCAPACITY BENEFIT

Year during episode	No of individuals	Percentage having frequent consultations (95% CI)	Odds ratio compared with period >5 years before claim (95% CI)
Years before start of claiming benefit:			
3	251	25 (20 to 31)	1.4 (0.85 to 2.2)
2	367	30 (25 to 35)	2.1 (1.4 to 3.2)
1	643	38 (34 to 42)	3.2 (2.2 to 4.6)
First year of claiming benefit	786	55 (52 to 59)	9.0 (6.2 to 13)

Adequacy of authors' replies to criticism raised in electronic letters to the editor: cohort study

Peter C Gøtzsche,¹ Tony Delamothe,² Fiona Godlee,² Andreas Lundh¹

EDITORIAL by Schriger and Altman

¹Nordic Cochrane Centre, Rigshospitalet and University of Copenhagen, Dept 3343, Blegdamsvej 9, DK-2100 Copenhagen, Denmark

²BMJ, BMA House, Tavistock Square, London

Correspondence to: P C Gøtzsche pcg@cochrane.dk

Cite this as: *BMJ* 2010;341:c3926 doi: 10.1136/bmj.c3926

This is a summary of a paper that was published on bmj.com as *BMJ* 2010;341:c3926

STUDY QUESTION

Is substantive criticism in letters to the editor, defined as a problem that could invalidate the research or reduce its reliability, adequately addressed by the authors?

SUMMARY ANSWER

Authors responded to substantive criticism in only about half of the cases.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS

Letters to the editor about research papers serve a useful role as post publication peer review and for the advancement of science. Editors should encourage authors to respond adequately—for example, by making a contract between the author and journal on acceptance of an article.

Participants and setting

We investigated consecutive research papers published in the *BMJ* between October 2005 and September 2007, and associated rapid responses.

Design, size, and duration

A cohort of 105 of 350 research papers published in the *BMJ* between October 2005 and September 2007 generated substantive criticism and were included in the main analyses.

Main results and the role of chance

A substantive criticism was raised against 105 of 350 (30%, 95% confidence interval 25% to 35%) included research papers, and of these the authors had responded to 47 (45%, 35% to 54%). The severity of the criticism was the same in those papers as in the 58 without authors' replies (mean score 2.2 in both groups, $P=0.72$). For the 47 criticisms with replies, no relation was found between the severity of criticism and the adequacy of the reply, neither as judged by the editors ($P=0.88$ and $P=0.95$, respectively), nor by the critics ($P=0.83$; response rate 85%). However, the critics were much more critical of the replies than the editors (average score 2.3 v 1.4, $P<0.001$).

Bias, confounding, and other reasons for caution

Our evaluations of the severity of criticisms and adequacy of replies were subjective.

Generalisability to other populations

Many more letters are published in the *BMJ* than in journals without online response systems. It is therefore not possible to extrapolate our results directly to journals that only publish letters in the print journal.

Study funding/potential competing interests

This study received no funding. The Nordic Cochrane Centre and *BMJ* provided in-house resources. TD and FG are editors of the *BMJ*.

Response on bmj.com

"The conclusion that authors should pay more attention to criticism is appropriate; the question is, why don't they? Perhaps once the paper is published, they have ticked that box and moved on to other work and have no sense of further commitment to engaging in a dialogue with others (critics or not). Perhaps some authors don't know how to engage in a critical discussion on their work, fearing it will undermine its integrity. In the end, it may be more about reputation than truth."

Michael Tremblay, healthcare adviser, Ashford, Kent

📌 To submit a rapid response, go to any article on bmj.com and select "Respond to this article"

EDITORS' AND CRITICS' OPINIONS OF ADEQUACY OF AUTHORS' REPLIES TO 40 PAPERS WITH SUBSTANTIVE CRITICISMS

Adequacy of replies to criticism	Editor 1*	Editor 2*	Critics*
Fully addressed (score 1)	29 (73)	27 (68)	6 (15)
Partly addressed (score 2)	7 (18)	7 (18)	16 (40)
Not addressed (score 3)	4 (10)	6 (15)	18 (45)

*Mean rating 1.4 for editors and 2.3 for critics ($P<0.001$).

What types of article does the *BMJ* consider?

We are delighted to receive articles for publication—from doctors and others—on the clinical, scientific, social, political, and economic factors affecting health. We give priority to articles that will help doctors to make better decisions. Please see our advice to authors at <http://resources.bmj.com/bmj/authors>, and if you would like to submit an article do so via our online editorial office at <http://submit.bmj.com>.

All original research articles are submitted, although we may invite submission (without promising acceptance) if we come across research being presented at conferences, if we see it in abstract form, or if the

authors make an inquiry about the suitability of their work before submission.

We are also pleased to consider submitted articles for sections which carry a mix of commissioned and submitted articles—editorials, analysis, clinical review, practice, fillers, and Career Focus. Please follow the specific advice on each of these article types (see <http://resources.bmj.com/bmj/authors/types-of-article>) before submitting your article. Some types of article—news, features, observations, head to head, views and reviews—are commissioned by the editors.

Cite this as: *BMJ* 2009;339:b3112