RESEARCH

The *BMJ* is an Open Access journal. We set no word limits on *BMJ* research articles, but they are abridged for print. The full text of each *BMJ* research article is freely available on bmj.com

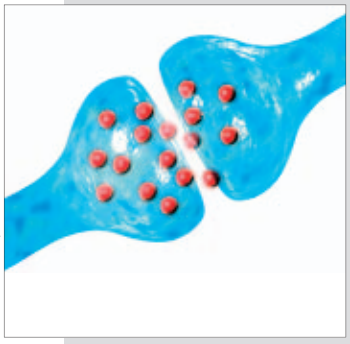
PHOTOS.COM

THIS WEEK'S RESEARCH QUESTIONS

Reboxetine and depression

Reboxetine is the first selective norepinephrine (noradrenaline) reuptake inhibitor approved for the treatment of depression in European countries. The drug works by binding to the norepinephrine transporter and blocking reuptake of extracellular norepinephrine, a catecholamine neurotransmitter. Selective serotonin reuptake inhibitors (SSRIs), the most commonly prescribed class of antidepressant, block the reuptake of the monoamine neurotransmitter serotonin instead.

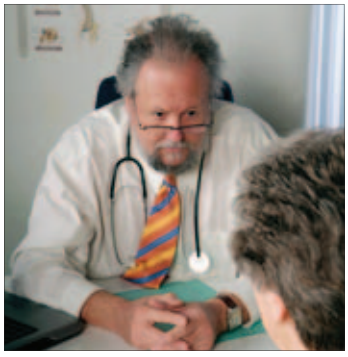
Although reboxetine has been claimed to be more effective than placebo and similarly effective to other antidepressants, Dirk Eyding and colleagues' systematic review and meta-analysis of published and unpublished trials has shown that reboxetine is an ineffective and potentially harmful antidepressant (p 816). Reboxetine was no better than placebo and inferior to SSRIs when it came to remission and response rates, whereas the drug was associated with higher rates of harmful outcomes than placebo and the SSRI fluoxetine.



LAGUNA DESIGN/SPL

Writing in an accompanying analysis article (p 809), the authors report on the difficulty they had obtaining unpublished studies and recommend "legal obligation for manufacturers to provide all requested data to health technology assessment bodies without commercial restrictions to publication."

Do we know whether patients are satisfied with English general practice?



CCSTUDIO/SPL

Under the Quality and Outcomes Framework in the United Kingdom patients are surveyed after seeing general practitioners, as one measure of practices' and doctors' performance. They're asked, among other things, "How long after your appointment time do you normally wait to be seen?" and "How do you rate your doctor's caring and concern for you?" Chris Salisbury and colleagues' secondary analysis of data from about 4500 patients at 27 English practices in 2005-6 found that reported experiences and satisfaction were related to patient sociodemographics, including age, sex, ethnicity, and employment status (p 820). But adjusting for patients' characteristics made little difference to the ranking of individual practices, and asking a global question about overall satisfaction wasn't much use.

Editorialist Jeannie Haggerty agrees with the authors' view that these measures need refining to include more information on real experiences rather than beliefs, but argues that they are not hopelessly flawed (p 790). When they do detect problems, says Professor Haggerty, these are real and important.

Reporting of euthanasia in Flanders, Belgium

The issue of euthanasia divides opinion among health professionals in the United Kingdom. Only last week doctors in the UK set up a group to campaign for assisted death for terminally ill, mentally competent adults (*BMJ* 2010;341:c5498), whereas many professional bodies, including the BMA, have policies against euthanasia.

In Belgium, euthanasia within strict guidelines has been legal since 2002. However, Tinne Smets and colleagues' research into reporting of cases to federal authorities suggests that doctors in Belgium are likewise not agreed on the issue (p 819). The authors found that only half (549 (52.8%)) of the 1049 cases of euthanasia in 2007 in Flanders, the Dutch speaking part of Belgium, were reported to the Federal Control and Evaluation Committee. Physicians were less likely to report a case if they did not perceive their act as euthanasia or if the time by which life was believed to have been shortened was less than a week.

The authors suggest that the current legal position in Belgium is insufficient and that policy needs strengthening to ensure physicians comply properly with the laws relating to a request for euthanasia, including their obligation to report.

ABRAHAM MENASHE/SPL

Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials

Dirk Eyding,[1] Monika Lelgemann,[2] Ulrich Grouven,[3 4] Martin Härter,[5] Mandy Kromp,[3] Thomas Kaiser,[3] Michaela F Kerekes,[3] Martin Gerken,[6] Beate Wieseler[3]

EDITORIAL by Godlee and Loder
ANALYSIS p 809

[1]German Cancer Society, Berlin, Germany (former affiliation: Institute for Quality and Efficiency in Health Care)
[2]Medical Advisory Service of Social Health Insurance, Essen, Germany
[3]Institute for Quality and Efficiency in Health Care, Cologne, Germany
[4]Hanover Medical School, Hanover, Germany
[5]University Clinic Eppendorf, Hamburg, Germany
[6]Health Technology Assessment Centre, University of Bremen, Bremen, Germany

Correspondence to: B Wieseler, Institute for Quality and Efficiency in Health Care, Dillenburger Strasse 27, 51105 Cologne, Germany
beate.wieseler@iqwig.de

Cite this as: *BMJ* 2010;341:c4737
doi: 10.1136/bmj.c4737

This is a summary of a paper that was published on bmj.com as *BMJ* 2010;341:c4737

STUDY QUESTION What are the benefits and harms of reboxetine in the acute treatment of major depressive disorder, and what is the impact of potential publication bias?

SUMMARY ANSWER Reboxetine is, overall, an ineffective and potentially harmful antidepressant, and the published evidence is substantially affected by publication bias.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS Reboxetine has been approved for the treatment of major depression in many European countries, but doubts have been raised about its efficacy and application for approval was rejected in the US. The present meta-analysis confirms the overall lack of efficacy and indicates potentially harmful effects of reboxetine, and provides a striking example of publication bias.

Selection criteria for studies

We considered double blind, randomised, controlled trials of acute treatment (six weeks or more) with reboxetine versus placebo or selective serotonin reuptake inhibitors (SSRIs) in adults with major depressive disorder. We searched bibliographic databases (Medline, Embase, PsycINFO, BIOSIS, and Cochrane Library), clinical trial registries, trial results databases, and regulatory authority websites up until February 2009, and requested unpublished data from the manufacturer of reboxetine (Pfizer).

Primary outcomes

The main outcomes analysed were remission and response rates (benefit outcomes), and rates of patients with at least one adverse event and withdrawals owing to adverse events (harm outcomes). The effect of publication bias was measured by comparing results of published trials with those from unpublished trials.

Main results and role of chance

The manufacturer did not provide unpublished data for the preliminary version of the health technology assessment report but supplied these data for the final analysis. We analysed 13 acute treatment trials that were placebo controlled, SSRI controlled, or both. These trials comprised 4098 patients; data on 74% of these patients were unpublished.

No significant differences in remission rates were shown in the trials that compared reboxetine with placebo (odds ratio 1.17, 95% confidence interval 0.91 to 1.51; P=0.216). Substantial heterogeneity (I^2=67.3%) was shown in the meta-analysis of response rates in patients receiving reboxetine and those receiving placebo, but a subsequent sensitivity analysis that excluded a small inpatient trial showed no significant difference in response rates (OR 1.24, 95% CI 0.98 to 1.56; P=0.071; I^2=42.1%). Reboxetine was inferior to SSRIs (fluoxetine, paroxetine, and citalopram) for remission rates (OR 0.80, 95% CI 0.67 to 0.96; P=0.015) and response rates (OR 0.80, 95% CI 0.67 to 0.95; P=0.01).

Reboxetine was inferior to placebo for both harm outcomes (P<0.001 for both), and to fluoxetine for withdrawals owing to adverse events (OR 1.79, 95% CI 1.06 to 3.05; P=0.031). Published data overestimated the benefit of reboxetine versus placebo by up to 115% and reboxetine versus SSRIs by up to 23%, and also underestimated harm.

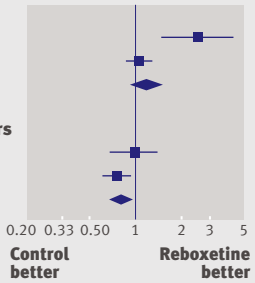
Bias, confounding, and other reasons for caution

The overall methodological quality of the trials was good bar one trial. At an outcome level, four trials showed a high risk of bias for response and remission rates; however, analyses excluding these outcomes did not alter the conclusions. The effect of publication bias on the results of the final analysis of all data was minor.

Study funding and potential competing interests This work was supported by the German Institute for Quality and Efficiency in Health Care (IQWiG). DE was employed by H Lundbeck A/S between January 2006 and April 2007. MH received remuneration from Boehringer Ingelheim and Lilly Pharma for three talks on depression guidelines in 2008.

META-ANALYSES OF REMISSION RATES IN PUBLISHED, UNPUBLISHED, AND ALL TRIALS

	Reboxetine (n/N)	Placebo or selective serotonin reuptake inhibitor (n/N)	Odds ratio (95% CI)	Odds ratio (95% CI)	Ratio of odds ratios; published:unpublished (95% CI)	Publication bias (%)
Reboxetine *v* placebo						
Published (1)	60/126	34/128		2.51 (1.49 to 4.25)		
Unpublished (6)	395/938	379/930		1.06 (0.88 to 1.28)	2.37 (1.36 to 4.13)	115
Total (7)	455/1064	413/1058		1.17 (0.91 to 1.51)		
Reboxetine *v* selective serotonin reuptake inhibitors						
Published (3)	123/247	132/260		0.98 (0.68 to 1.40)		
Unpublished (5)	372/824	432/833		0.75 (0.61 to 0.94)	1.31 (0.86 to 1.99)	23
Total (8)	495/1071	564/1093		0.80 (0.67 to 0.96)		



Publication bias (right column) is presented as the ratio of odds ratios of published results versus overall results. The extent of publication bias is expressed as percentage change between the analysis of published trials only and the analysis of all trials (that is, publication bias=100 x ($OR_{published\ data}/OR_{total\ data}-1$)).

Association of cerebral palsy with Apgar score in low and normal birthweight infants: population based cohort study

Kari Kveim Lie,[1] Else-Karin Grøholt,[1] Anne Eskild[2 3]

EDITORIAL by Paneth

[1]Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway
[2]Department of Obstetrics and Gynaecology and Medical Faculty Division, Akershus University Hospital, Nordbyhagen, Norway
[3]Division of Mental Health, Norwegian Institute of Public Health, Oslo, Norway
Correspondence to: K Kveim Lie, Division of Epidemiology, Norwegian Institute of Public Health, PO Box 4404, Nydalen, 0403 Oslo, Norway
kari.kveim.lie@fhi.no

Cite this as: *BMJ* 2010;341:c4990
doi: 10.1136/bmj.c4990

This is a summary of a paper that was published on bmj.com article as *BMJ* 2010;341:c4990

STUDY QUESTION How is Apgar score in newborn infants associated with cerebral palsy in childhood?

SUMMARY ANSWER Low Apgar score 5 minutes after birth is strongly associated with cerebral palsy in children with normal birth weight and more modestly associated in children with low birth weight.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS Previous studies have shown that low Apgar score is associated with cerebral palsy in children born to term, but a possible association in low birthweight children has been debated. This paper shows that low Apgar score is also significantly associated with cerebral palsy in low birthweight children.

Participants and setting

The Medical Birth Registry of Norway was used to identify all babies born between 1986 and 1995. Information on the diagnosis of cerebral palsy was obtained by linkage to the Registry of Cerebral Palsy in Children born 1986-95, which is based on discharge diagnoses from all paediatric departments in Norway.

Design, size, and duration

We undertook a population based cohort study of 543 064 singletons without malformations who survived the first year of life.

Main results and the role of chance

A total of 988 children (0.1%) were diagnosed with cerebral palsy. In total, 11% (39/369) of the children with Apgar score of less than 3 at birth were diagnosed with cerebral palsy, compared with only 0.1% (162/179 515) of the children with Apgar score of 10 (odds ratio (OR) 53, 95% CI 35 to 80 after adjustment for birth weight).

In children with birth weight of less than 1500 g, 17% of those with Apgar score of less than 4 developed cerebral palsy compared with 4% of those with an Apgar score of more than 8 (odds ratio 5, 95% CI 2 to 9). In children with birth weight of 2500 g or more, 10% of those with an Apgar score of less than 4 developed cerebral palsy compared with only 0.1% of those with Apgar score more than 8 (odds ratio 125, 95% CI 91 to 170).

Low Apgar score was strongly associated with each of the three subgroups of spastic cerebral palsy, although the association was strongest for quadriplegia (OR 137 for Apgar score <4 *v* Apgar score >8, 95% CI 77 to 244 after adjustment for birth weight).

Bias, confounding, and other reasons for caution

Information on diagnoses of cerebral palsy was obtained from all paediatric departments in Norway. We believe, therefore, that most children in the birth cohort with a diagnosis of cerebral palsy were identified and included in the study.

Generalisability to other populations

Low Apgar score might reflect damage of the central nervous system that has occurred before as well as during birth. The proportion of cases with cerebral palsy that could be attributed to such damage, either before or during birth, may depend on the level of obstetric care.

Knowledge of increased prevalence of cerebral palsy in children with low Apgar score may be important in clinical follow-up of children. It should, however, be noted that almost 90% of children with Apgar score less than 4 did not develop cerebral palsy.

Study funding/potential competing interests

The study was funded by the Norwegian Foundation for Health and Rehabilitation. The authors declare no conflicts of interest.

PREVALENCE AND ODDS OF CEREBRAL PALSY WITHIN BIRTHWEIGHT GROUPS ACCORDING TO APGAR SCORE

Apgar score	Proportion with cerebral palsy (n/N (%))	Odds ratio (95% CI)
Birth weight <1500 g (n=2629)		
<4	12/71 (16.9)	4.6 (2.3 to 9.0)
4-6	35/349 (10.0)	2.5 (1.6 to 3.9)
7-8	90/887 (10.1)	2.6 (1.8 to 3.6)
>8	56/1322 (4.2)	Reference
Birth weight 1500-2499 g (n=15 357)		
<4	8/71 (11.3)	15.7 (7.3 to 33.6)
4-6	21/449 (4.7)	6.1 (3.8 to 9.8)
7-8	59/2012 (2.9)	3.7 (2.7 to 5.2)
>8	103/12 825 (0.8)	Reference
Birth weight ≥2500 g (n=525 078)		
<4	48/495 (9.7)	124.5 (91.1 to 170.1)
4-6	64/2761 (2.3)	27.5 (21.1 to 35.9)
7-8	58/18 353 (0.3)	3.7 (2.8 to 4.8)
>8	434/503 469 (0.1)	Reference

Variation in rates of caesarean section among English NHS trusts after accounting for maternal and clinical risk: cross sectional study

Fiona Bragg,[1 2] David A Cromwell,[1 4] Leroy C Edozien,[3] Ipek Gurol-Urganci,[1 4] Tahir A Mahmood,[4] Allan Templeton,[4] Jan H van der Meulen[1 4]

EDITORIAL by Knight and Sullivan

[1]London School of Hygiene and Tropical Medicine, London, UK
[2]NHS City and Hackney, London, UK
[3]Maternal and Fetal Health Research, St Mary's Hospital, Manchester, UK
[4]Royal College of Obstetricians and Gynaecologists, London, UK
Correspondence to: David Cromwell, Health Services Research Unit, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK
david.cromwell@lshtm.ac.uk

Cite this as: *BMJ* 2010;341:c5065
doi: 10.1136/bmj.c5065

This is a summary of a paper that was published on bmj.com as *BMJ* 2010;341:c5065

STUDY QUESTION

Can the variation in the rates of caesarean section among English NHS trusts be explained by maternal characteristics and clinical risk factors?

SUMMARY ANSWER

Unadjusted rates of caesarean section among 146 NHS trusts ranged from 13.6% to 31.9% and adjusted rates ranged from 14.9% to 32.1%; however, only rates of emergency caesarean section varied notably between trusts.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS

Publication of the rates of caesarean section for English NHS trusts in 2008-9 led to debate about the causes of the large variation observed and the higher rates in the south of England (north-south divide). This study showed that differences in hospitals' patients did not explain the between trust variation but eliminated the north-south divide. Attempts to reduce the variation should examine issues linked to emergency caesarean sections.

Participants and setting

This study included women aged between 15 and 44 years who delivered a singleton baby between 1 January and 31 December 2008 in all English NHS trusts with more than 1000 deliveries that year.

Design, size, and duration

Hospital episode statistics data were used to analyse patterns of caesarean deliveries (elective and emergency) in 146 English NHS trusts. We used a multiple logistic regression model to estimate the likelihood of women having a caesarean section given their maternal characteristics (age, ethnicity, parity, socioeconomic deprivation) and clinical risk factors (previous caesarean section, breech presentation, fetal distress).

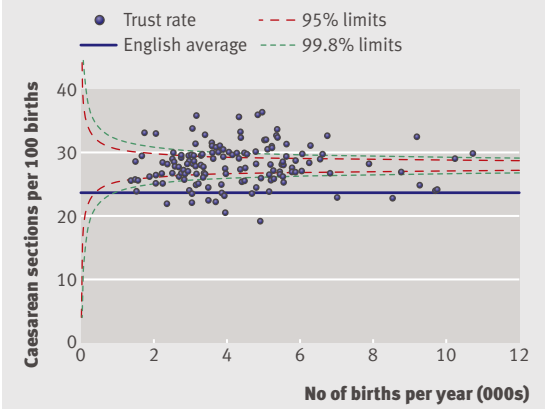
Primary outcomes

The primary outcomes were unadjusted and adjusted rates of caesarean section per 100 births (live or stillborn).

Main results and the role of chance

Among 620 604 births, 147 726 (23.8%) were delivered by caesarean section. Women were more likely to have undergone a caesarean section if they had had a caesarean previously (70.8%) or had a baby with breech presentation (89.8%). Unadjusted rates of caesarean section among the NHS trusts ranged from 13.6% to 31.9%.

ADJUSTED RATES OF CAESAREAN SECTION IN 2008



Trusts differed in their patient populations, but adjusted rates still ranged from 14.9% to 32.1% (figure). Rates of emergency caesarean section varied between trusts more than rates of elective caesarean section.

The variation in the overall rates mainly arose from differences in the rates of emergency caesarean section: the 10th and 90th percentiles of the adjusted rates of emergency caesarean section were 10.7% and 18.9%, whereas for elective caesarean section these percentiles were 7.8% and 11.2%.

Bias, confounding, and other reasons for caution

The results could be influenced by inaccuracies in the coding of delivery method. Broad categories of caesarean section were used to improve reliability, and previous studies have reported good agreement between routine databases like hospital episode statistics and medical notes for caesarean sections.

The hospital episode statistics contained incomplete and inaccurate information on factors known to influence mode of delivery, which limited the selection of variables for inclusion in the risk adjustment model. Some residual confounding is likely, but the discrimination of the logistic regression model was very good.

Generalisability to other populations

The study's findings apply to the United Kingdom only. Variation in rates of caesarean section affects other countries, but the reasons for it may differ.

Study funding and potential competing interests

The study did not receive external funding. JHvdM received a national public health career scientist award from the Department of Health and NHS research and development programme.

Response on bmj.com

"Previously, concerns have been raised about a blanket approach to collecting data on women's weight . . . These concerns need to be resolved without further delay, and progress made, so that women can receive counselling about weight gain, nutrition, and food choices, while at the same time being sensitively advised of some of the possible increased risks."
Pauline M Hull, electivecesarean.com

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

Reporting of euthanasia in medical practice in Flanders, Belgium: cross sectional analysis of reported and unreported cases

Tinne Smets,[1] Johan Bilsen,[1] Joachim Cohen,[1] Mette L Rurup,[2] Freddy Mortier,[3] Luc Deliens[1 2]

[1]End of Life Care Research Group, Vrije Universiteit Brussel, Brussels, Belgium
[2]Department of Public and Occupational Health, EMGO Institute for Health and Care Research, Expertise Center for Palliative Care, VU University Medical Centre, Amsterdam, Netherlands
[3]Bioethics Institute Ghent, Ghent University, Ghent, Belgium

Correspondence to: T Smets, End of Life Care Research Group, Faculty of Medicine and Pharmacy, Laarbeeklaan 103, 1090 Brussels, Belgium **tinne.smets@vub.ac.be**

Cite this as: *BMJ* 2010;341:c5174
doi: 10.1136/bmj.c5174

This is a summary of a paper that was published on bmj.com as *BMJ* 2010;341:c5174

STUDY QUESTION What is the estimated rate at which physicians in Flanders, Belgium, report euthanasia cases to the Federal Control and Evaluation Committee?

SUMMARY ANSWER The reporting rate for euthanasia in Flanders, Belgium, in 2007 is estimated at 52.8%.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS To provide societal control over the practice of euthanasia, physicians in Belgium are required by law to report each case to the Federal Control and Evaluation Committee. Only half of all cases of euthanasia in Flanders in 2007 were being reported; unreported cases were often not perceived as euthanasia by the physician and were generally dealt with less carefully than reported cases.

Participants and setting

A stratified at random sample was drawn of people who died between 1 June 2007 and 30 November 2007 in Flanders, Belgium. The certifying physician of each death was sent a questionnaire on end of life decision making in the death concerned.

Design

Cross sectional analysis.

Primary outcomes

The rate of euthanasia cases reported to the Federal Control and Evaluation Committee, physicians' reasons for not reporting cases of euthanasia, the relation between reporting and non-reporting and the characteristics of the physician and patient, the time by which life was shortened according to the physician, the labelling of the end of life decision by the physician involved, and differences in characteristics of due care between reported and unreported euthanasia cases.

REPORTING RATES FOR EUTHANASIA IN FLANDERS, BELGIUM, IN 2007

	Number of cases	Rate
Estimated number of cases of euthanasia	137	—
Estimated number of reported cases of euthanasia	549	—
Estimated weighted total number of cases of euthanasia*	1040	1.9% (1.6% to 2.3%)†
Overall reporting rate for euthanasia‡		52.8% (43.9% to 60.5%)†
Reporting rates for euthanasia according to drug use‡§		
Recommended drugs¶	70	92.9% (84.3% to 96.5%)
Non-recommended drugs**	61	4.8% (1.1% to 16.9%)

*The estimated total rate of euthanasia was calculated by weighting for stratification and for patient and mortality characteristics of all deaths in 2007. The original number of euthanasia cases in the sample was 137. One case was missing data on the variable "reporting of end of life decision."
†Percentage of all deaths in Flanders, Belgium, 2007.
‡Weighted percentage.
§Five "missings" on the variable "drugs used for euthanasia."
¶Barbiturates, neuromuscular relaxants, or both.
**Opioids, benzodiazepines, or drugs other than barbiturates or neuromuscular relaxants.

Main results and the role of chance

The survey response rate was 58.4% (3623/6202 eligible cases). The estimated total number of cases of euthanasia in Flanders in 2007 was 1040 (95% CI 970 to 1109); thus the incidence of euthanasia was estimated as 1.9% of all deaths (95% CI 1.6% to 2.3%). Approximately half (549/1040 (52.8%, 95% CI 43.9% to 60.5%)) of all estimated cases of euthanasia were reported to the Federal Control and Evaluation Committee. Physicians who perceived their case as euthanasia reported it in 93.1% (67/72) of cases. Cases of euthanasia were reported less often when the time by which life was shortened was less than one week compared with when the perceived life shortening was greater (37.3% *v* 74.1%; P<0.001). Unreported cases were generally dealt with less carefully than reported cases: a written request for euthanasia was more often absent (87.7% *v* 17.6% verbal request only; P<0.001), other physicians and care givers specialised in palliative care were consulted less often (54.6% *v* 97.5%; P<0.001; 33.0% *v* 63.9%; P<0.001), the life ending act was more often performed with opioids or sedatives (92.1% *v* 4.4%; P<0.001), and the drugs were more often administered by a nurse (41.3% *v* 0.0%; P<0.001).

Bias, confounding and other reasons for caution

The response rate was only 58%; thus a non-response bias cannot be completely excluded. Furthermore, the study is based on self reporting by physicians; as such, social desirability bias and the possibility that physicians did not remember all aspects of a case cannot be excluded.

Generalisability to other populations

Our findings cannot be extrapolated to the French speaking part of Belgium, in particular because research has shown that end of life practices differ in the French speaking and the Flemish speaking regions and because there may be a difference in willingness to report cases of euthanasia owing to cultural differences. Our results on reported and unreported cases are very similar to those found in the Netherlands, but the current reporting rate in Flanders is much lower than that in the Netherlands. Nevertheless, the study offers valuable data driven information that can inform the debates about the legalisation of euthanasia that are currently going on in the United Kingdom and in many other countries.

Study funding/potential competing interests

This study is supported by a grant from the Institute for the Promotion of Innovation by Science and Technology in Flanders (Instituut voor de aanmoediging van Innovatie door Wetenschap en Technologie in Vlaanderen; SBO IWT nr. 050158). The funders had no role in study design; in the collection, analysis and interpretation of the data; in the writing of the manuscript; or in the decision to submit for publication. The authors declare no competing interests.

Patients' experience and satisfaction in primary care: secondary analysis using multilevel modelling

Chris Salisbury,[1] Marc Wallace,[2] Alan A Montgomery[1]

EDITORIAL by Haggerty

[1]University of Bristol, Bristol BS8 2AA, UK
[2]Royal Berkshire Hospital, Reading RG1 5AN, UK
Correspondence to: C Salisbury
c.salisbury@bristol.ac.uk

Cite this as: *BMJ* 2010;341:c5004
doi: 10.1136/bmj.c5004

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2010;341:c5004

STUDY QUESTION Do responses to questions in surveys of patients that purport to assess the performance of general practices or doctors reflect differences between practices, doctors, or patients themselves?

SUMMARY ANSWER

Patients' reports of their experiences, as well as satisfaction, are systematically related to their own characteristics such as age and sex, but the effect of adjusting practices' scores for the characteristics of their patients is small.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS

Surveys of patients are used to assess the performance of doctors and practices, but few studies have explored the extent to which variation in reported satisfaction and experience is due to differences between practices, doctors, or patients themselves. Questions about patients' satisfaction discriminate poorly between practices and doctors, but questions about specific experiences are more discriminatory; adjusting for patients' characteristics makes little difference to practices' performance scores.

Participants and setting

The study was based on 4573 patients who had consulted 150 different doctors from 27 general practices in England.

Design

We did a secondary analysis of data from a study of access systems in general practice. Patients completed a questionnaire survey after their consultation. We combined data from their responses with information about the practices and the doctors consulted. We used multilevel modelling to explore whether variation in patients' responses to questions about their satisfaction and their experiences reflected differences between practices or between doctors and to what extent variation was explained by the patients' own characteristics.

Primary outcomes

The main outcomes were overall satisfaction, experience of length of wait for an appointment, perceived access to care, and satisfaction with communication skills.

Main results and the role of chance

The survey had an 84.0% response rate. The experience based measure of how long the patient waited for an appointment was more discriminating between practices (differences between practices accounted for 20.2% of the variance) than was the overall satisfaction measure (practices accounted for 4.6% of variance). Only 6.3% of the variance in responses to questions about doctors' communication skills was due to differences between doctors. Patients' reports of their wait for an appointment were related to practices' organisational characteristics, whereas their satisfaction with doctors' communication skills was related to the doctors' age and where they qualified. Patients' reports of their experiences, as well as their satisfaction, were related to characteristics such as patients' age, sex, ethnicity, and housing and employment status. However, adjusting for patients' characteristics made little difference to the ranking of individual practices.

Bias, confounding, and other reasons for caution

Patients were included in the study because they attended a consultation, so the findings will be weighted towards those who attend most often. The number of practices included and the number of patients per doctor were relatively small. Larger samples would have provided greater ability to discriminate between doctors and practices, but the key messages about the relative discriminatory power of different types of questions would remain. The exclusion of patients with missing data may introduce bias if data are not missing completely at random.

Generalisability to other populations

Practices were included in this study because they agreed to participate in a research study, so findings may not be generalisable to all practices. The finding that adjustment for patients' characteristics made little difference to practices' scores may not apply to practices with very atypical populations.

Study funding/potential competing interests

The study of access systems was funded by the NHS Research and Development Programme on Service and Delivery Organisation. This secondary analysis was unfunded.

PERCENTAGE OF VARIANCE EXPLAINED AT LEVEL OF PRACTICE, DOCTOR, AND PATIENT FOR EACH OUTCOME

	Estimated variance (95% CI) as percentage of total variance in outcome		
Outcome	Between practice	Between doctor	Patient plus random
Overall satisfaction (single question) (n=4414)	4.6 (1.6 to 7.6)	1.5 (0.4 to 2.6)	93.9 (89.8 to 98.0)
Wait for appointment (single question) (n=4058)	20.2 (9.1 to 31.3)	0.8 (0 to 1.7)*	79.1 (75.5 to 82.6)
Access scale† (n=4517)	14.9 (6.4 to 23.3)	1.5 (0.4 to 2.6)	83.7 (80.2 to 87.2)
Communication scale† (n=4423)	1.2 (0 to 3.0)*	6.3 (3.8 to 8.9)	92.4 (88.5 to 96.4)

*Lower confidence limit restricted to 0.
†From general practice assessment questionnaire.