Open Access: free full text available online
For the full versions of these articles see bmj.com

Effectiveness of policy to provide breastfeeding groups (BIG) for pregnant and breastfeeding mothers in primary care: cluster randomised controlled trial

Pat Hoddinott,[1] Jane Britten,[1] Gordon J Prescott,[2] David Tappin,[3] Anne Ludbrook,[4] David J Godden[1]

EDITORIAL by Cattaneo
RESEARCH p 392

[1]Centre for Rural Health, University of Aberdeen, Centre for Health Science, Inverness IV2 3JH
[2]Department of Public Health, University of Aberdeen, Aberdeen AB25 2ZD
[3]Paediatric Epidemiology and Community Health Unit, Section of Child Health, University of Glasgow, Royal Hospital for Sick Children, Glasgow G3 8SJ
[4]Health Economics Research Unit, University of Aberdeen, Aberdeen
Correspondence to: P Hoddinott
p.hoddinott@abdn.ac.uk

Cite this as: *BMJ* 2009;338:a3026
doi:10.1136/bmj.a3026

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;338:a3026

ABSTRACT

Objective To assess the clinical effectiveness and cost effectiveness of a policy to provide breastfeeding groups for pregnant and breastfeeding women.

Design Cluster randomised controlled trial with prospective mixed method embedded case studies to evaluate implementation processes.

Setting Primary care in Scotland.

Participants Pregnant women, breastfeeding mothers, and babies registered with 14 of 66 eligible clusters of general practices (localities) in Scotland that routinely collect breastfeeding outcome data.

Intervention Localities set up new breastfeeding groups to provide population coverage; control localities did not change group activity.

Main outcome measures Primary outcome: any breast feeding at 6-8 weeks from routinely collected data for two pre-trial years and two trial years. Secondary outcomes: any breast feeding at birth, 5-7 days, and 8-9 months; maternal satisfaction.

Results Between 1 February 2005 and 31 January 2007, 9747 birth records existed for intervention localities and 9111 for control localities. The number of breastfeeding groups increased from 10 to 27 in intervention localities, where 1310 women attended, and remained at 10 groups in control localities. No significant differences in breastfeeding outcomes were found. Any breast feeding at 6-8 weeks declined from 27% to 26% in intervention localities and increased from 29% to 30% in control localities (P=0.08, adjusted for pre-trial rate). Any breast feeding at 6-8 weeks increased from 38% to 39% in localities not participating in the trial. Women who attended breastfeeding groups were older (P<0.001) than women initiating breast feeding who did not attend and had higher income (P=0.02) than women in the control localities who attended postnatal groups. The locality cost was £13 400 (€14 410; $20 144) a year.

Conclusion A policy for providing breastfeeding groups in relatively deprived areas of Scotland did not improve breastfeeding rates at 6-8 weeks. The costs of running groups would be similar to the costs of visiting women at home.

Trial registration Current Controlled Trials ISRCTN44857041.

INTRODUCTION

Evidence for short term and long term health benefits of breast feeding for mother and child is increasing.[1 2] Scotland has among the lowest breastfeeding rates in the developed world; only 44% of babies received any breast milk at 6 weeks in 2005.[3]

Two Cochrane reviews have summarised the evidence for interventions that increase the prevalence of breast feeding.[4 5] One to one professional or lay support increases the duration of any breast feeding up to 6 months, with a greater effect for exclusive breast feeding.[5] The effectiveness of health service provided breastfeeding groups is unknown. Our aim was to evaluate the clinical effectiveness and cost effectiveness of a policy for providing breastfeeding groups for pregnant women and breastfeeding mothers.

METHODS

Participants—All 66 clusters of general practices (localities) that routinely collected breastfeeding data through the Child Health Surveillance Programme (CHSP) of the National Health Service (NHS) Scotland from October 2002 were eligible to participate. We recruited 14 localities (see bmj.com).

Interventions—Locality staff implemented the policy largely within existing resources. The figure shows a time line describing the complex intervention and associated research activity,[6] and the box describes the policy.

Outcomes—Our primary outcome was the number of babies receiving any breast milk at 6-8 weeks. Secondary outcomes were the number of babies receiving any breast milk at birth, 5-7 days, and 8-9 months, as well as maternal satisfaction and social support. We collected quantitative and qualitative data (figure) to build case studies at two levels: the case defined as the locality (intervention and control) to investigate differences in implementation of the intervention and in routine care between all areas and the case defined as the group to evaluate group styles and processes. We did four structured telephone surveys with 45 key informants from 14 localities to monitor breastfeeding group activity, other breastfeeding

initiatives, and changes in maternity and child health services during the intervention.

Costs and benefits—To evaluate the costs and benefits to women, we asked group facilitators to give questionnaires to all women attending breastfeeding groups (intervention) and postnatal groups (control) during three separate one week periods throughout the trial, covering different times of year, and spaced out to minimise the risk of including women more than once. The main NHS cost was staff time.

Statistical methods—We analysed all outcomes at cluster level on an intention to treat principle, blinded to allocation. We analysed the difference in breastfeeding rates between intervention and control localities by using analysis of covariance, with pre-intervention breastfeeding rates as a covariate (see bmj.com).

Breastfeeding group policy

In localities

Amount of intervention to be achieved
We asked each intervention locality to at least double the amount of breastfeeding group activity, set up a minimum of two new breastfeeding groups, and ensure that all main population centres had access to a breastfeeding group

Resources provided
Each locality appointed a local investigator who was a health visitor with an interest in breast feeding. Resources were provided to protect her time for one day a week for the four month set-up period and half a day a week for the two intervention years. Each locality nominated a principal investigator (no additional funding provided), who was a public health practitioner or health promotion officer, to oversee trial implementation. Midwives and health visitors facilitated groups as part of routine antenatal and postnatal care, supported by local volunteers, other interested health service staff, and students. We provided written information for women and posters

Implementation of the policy
We asked each locality to set up steering group meetings every six to eight weeks for group facilitators and for participant, voluntary sector, and relevant locality stakeholder representatives. The aim was for steering groups to capture the multidisciplinary sharing of experiences, support, and reflective practice, which was an important component of the preliminary action research study.[7-9] We asked steering groups to reflect on what was working well, what was not working so well, and what could be changed within the scope of the trial protocol and to review any changes made

Pre-trial training
We offered two half day training seminars for 20 health professionals in each locality: one on group facilitation skills led by an external consultant and one on trial conduct, protocol, and data collection. We provided a written training pack and a password protected website with access to all training materials

In individual breastfeeding groups

The following six aspects of breastfeeding groups were fixed across all seven intervention localities. These characteristics were derived from the successful preliminary action research study.[7-9]

- Weekly group meetings
- Women only
- A health professional group facilitator must be present
- Pregnant women and breastfeeding women to be invited to attend; implicit in this is the need for health visitors and midwives to work together to recruit women to attend groups
- At least 50% of the group meeting time to be social and interactive
- A woman centred approach to group timing, content, and structure based on the women's needs

RESULTS

At baseline, intervention localities had fewer general practices classified as rural, had fewer maternity units, and were slightly less deprived than control localities but were otherwise comparable. Intervention localities increased breastfeeding groups from 10 to 27, and control localities remained unchanged with 10 groups.

Feeding outcomes—The table reports the proportion of babies with valid records receiving any breast milk. We found no significant differences in rates of any breast feeding at 6-8 weeks, birth, 5-7 days, or 8-9 months.

Intervention women were three months older than controls ($P<0.001$) and slightly less deprived both before and after the intervention (χ^2 tests $P<0.001$). We found small increases in maternal deprivation from before to after the intervention for intervention localities ($P=0.021$), control localities ($P=0.056$), and external control localities ($P<0.001$).

Baby-friendly hospital initiative status—Breastfeeding data by baby-friendly hospital initiative status is available on www.abdn.ac.uk/crh/big.shtml. In the seven localities where the main hospital received the award or a certificate of commitment towards the award in 2005, initiation of breast feeding increased or remained the same. For the six localities that gained the award before the pre-trial period, we found no clear pattern.

Satisfaction and support—We found no significant differences in maternal satisfaction or social support between intervention and control localities (see bmj.com).

Group characteristics—From 2195 group diaries, 2007 weekly group meetings took place, 188 were cancelled, and 1310 unique women (pregnant or breast feeding) attended. Median group size was four (interquartile range 2-6) women, excluding group facilitators, children, female relatives, non-pregnant friends, and students. Significantly more women in intervention than control localities knew that breastfeeding groups (antenatal and postnatal) were available and attended postnatal breastfeeding groups (data available on www.abdn.ac.uk/crh/big.shtml). Only 229/799 (28.7%) women who initiated breast feeding reported attending antenatal breastfeeding groups in intervention localities compared with 99/416 (23.8%) respondents in control localities ($P=0.093$). Each intervention locality held a median of 46 (25-64) group meetings per 100 women initiating breast feeding per year, and a median of 23 (19-30) women attended per 100 women initiating breast feeding per year. Individual locality breastfeeding outcomes could not be explained by group attendance or the number of group meetings held.

Costs and benefits—Staff time varied considerably between groups owing to differences in the number of

	Policy to provide breastfeeding groups (intervention)	Usual care with no new breastfeeding group activity (control)	Non-participating localities collecting outcome data (external control)
2 years pre-intervention	Measurement of breastfeeding outcomes (1 October 2002 to 30 September 2004)		
Randomisation			
4 months preparation	a		
Baseline	b c d e	e	
2 years from baseline	b c d e f g h i j k	e f j	
2 years from baseline	Measurement of breastfeeding outcomes (1 February 2005 to 31 January 2007)		

a	Training day and information packs for each locality for <20 health visitors, midwives, or others likely to be involved in groups
b	Double pre-trial breastfeeding group activity, cover main populations in locality, and set up minimum of two new groups for two years
c	Run groups according to the protocol
d	Set up and hold locality steering group meetings every 6-8 weeks for group facilitators, participants, voluntary sector, and relevant locality stakeholder representatives to reflect on policy implementation
e	Structured telephone interviews every six months with key informants to map all breastfeeding related activity in each locality
f	Health visitors distribute (months 6-24) questionnaire at routine 6-8 week baby check appointment to women who initiated breastfeeding (put baby to breast at least once after birth). Questionnaire includes maternal satisfaction with breastfeeding experience; social support; attendance at general birth related and specific breastfeeding groups, classes, or workshops in pregnancy and after birth, with free text question about experiences attending groups
g	Group meeting attendance registers, diaries, and first time participant characteristics questionnaire
h	Structured group observations (n=17)
i	Qualitative interviews (n=105 participants in 126 interviews). Focus groups (n=13) with health professionals and peer supporters. Individual or pair interviews (face to face (n=41) or telephone (n=27)) with group facilitators, minimally involved health professionals, nurse managers, women group attenders, and non-attenders
j	Questionnaire to breastfeeding group (intervention) and postnatal group (control) participants to determine costs including travel, childcare, time, and lost income, and benefits and value of groups (willingness to pay questions)
k	Group facilitator workload survey over one week to assess staff costs

Timing and characteristics of intervention

staff attending and the length of group meetings. The average cost per locality per year was about £13 400 (€14 410; $20 144), including travel time (2005/6 prices). The cost equates to £143 per unique woman attending (n=1310). Completed costs and benefits questionnaires were returned by 175 women attending intervention breastfeeding groups (a minimum response rate of 53%) and by 156 women attending control postnatal groups (a minimum response rate of 27%). Little difference existed in time and travel costs between attending a breastfeeding group or a postnatal group; most women travelled less than 5 miles, and most journeys took less than 10 minutes. Mean willingness to pay to attend the groups was £2.42 for the intervention group and £2.54 for the control group; the difference between the groups was not statistically significant.

DISCUSSION

This study shows that a policy to provide breastfeeding groups has no impact on rates of breast feeding and highlights several challenges in the pragmatic implementation of health promotion initiatives. The embedded mixed method case study design provides insight into why the preliminary study was effective and the trial was not.[7] In the preliminary study, attendance by pregnant women was six times higher and attendance by any woman (pregnant or breast feeding) was four times higher, more midwives facilitated groups, and three out of five groups took place in community maternity units compared with only one (the best attended) of 27 groups in the trial. More centralised maternity services, reorganisation, and workforce pressures contributed to these differences.[10] The preliminary study used action

Any breast feeding (exclusive or partial) as a proportion of those with known feeding behaviour

Primary and secondary outcomes	Mean (SD) breastfeeding rate		Mean difference between groups (95% CI for difference)*	P value for difference*	Non-participating localities† (overall proportion)
	Intervention	Control			
At 6-8 weeks					
Pre-intervention	0.27 (0.03)	0.29 (0.08)			0.38
Post-intervention	0.26 (0.03)	0.30 (0.07)	−0.017 (−0.036 to 0.002)	0.08	0.39
At birth					
Pre-intervention	0.50 (0.05)	0.51 (0.10)			0.59
Post-intervention	0.51 (0.06)	0.53 (0.09)	−0.009 (−0.045 to 0.027)	0.58	0.60
At 5-7 days					
Pre-intervention	0.43 (0.04)	0.46 (0.09)			0.55
Post-intervention	0.42 (0.04)	0.45 (0.09)	−0.003 (−0.032 to 0.027)	0.84	0.55
At 8-9 months					
Pre-intervention	0.22 (0.03)	0.23 (0.04)			0.28
Post-intervention	0.21 (0.05)	0.20 (0.06)	0.007 (−0.056 to 0.070)	0.82	0.25

*Adjusted for pre-intervention rate.
†Child Health Surveillance Programme data are for all NHS board areas that collect such data, excluding trial localities; Guthrie data are for all Scotland, excluding trial localities.

research methods,[11] compared with a distant research team running a trial, and partnership working between midwives, health visitors, and women was less evident in the trial. Involvement of midwives is crucial to recruit pregnant women to groups, as health visitors' first contact is usually 10-14 days after birth, when 17% of women have stopped breast feeding.[3] The discontinuity of professional care during women's breastfeeding journey from pregnancy to weaning warrants review, as maternity care has changed considerably with shorter hospital stays.

Limitations

Participating localities were more deprived and had lower baseline breastfeeding rates than the Scottish average. However, the large numbers of women included may have inflated the importance of the small changes in maternal age and deprivation seen. Our findings may not be generalisable to less deprived populations or countries with higher breastfeeding rates, where breast feeding is more socially acceptable and sustaining groups might be easier. Qualitative case study data (to be reported separately) suggest that operational factors, particularly socio-geographic characteristics of localities including deprivation, staff resources, the amount of organisational change, and the style of management and leadership, all affected implementation of the trial and seem to explain why breastfeeding rates declined in three intervention localities.

During the intervention, more babies were born in hospitals that achieved the baby-friendly hospital initiative award or the certificate of commitment in control localities, and initiation of breast feeding increased in these localities. In retrospect, some people might argue that we could have matched by baby-friendly hospital status, given the evidence that it increases duration of breast feeding.[12] However, our data support the view that in the UK the initiative does not have an effect beyond initiation of breast feeding and that earlier research showing an increase in breast feeding at 5-7 days might depend on the amount of time since the award was achieved.[13 14] The changes in status therefore seem unlikely to have outweighed the effects of the intervention, but we cannot rule this out.

Attending groups and developing social networks may have other short term and long term benefits that we did not measure. Few conclusions can be drawn from our questionnaire derived secondary outcomes owing to the low response rate. Reported attendance rates were higher at general antenatal groups than at breastfeeding groups in both trial arms, and we would recommend review of the current practice of providing separate antenatal breastfeeding education.

Conclusion

If the NHS aims to increase breastfeeding rates, a policy to provide breastfeeding groups, despite being a low cost option that women attending find helpful, is nevertheless ineffective in the current organisational context. Resources may be better directed to the first two weeks after birth, when the highest proportion of women are stopping breast feeding.

We thank all the managers, health professionals, volunteers, and women who participated in this trial and made it possible. We thank Lee Dowie for trial administration, Fiona Ryan for secretarial support, Roisin Pill (Cardiff University) for qualitative research consultancy, Marion Campbell

WHAT IS ALREADY KNOWN ON THIS TOPIC

Additional professional and lay support increases the duration and exclusivity of breast feeding

Little is known about the effectiveness of a group setting for breastfeeding interventions

Older and higher income women are more likely to initiate breast feeding

WHAT THIS STUDY ADDS

Breastfeeding support groups, facilitated by health professionals, for pregnant and breastfeeding women did not improve breastfeeding rates in the first six to eight weeks after birth

Of women initiating breast feeding, older women were more likely to attend groups; women attending breastfeeding groups had a higher income than those attending postnatal groups

The costs of providing groups are similar to the costs of home visits by health visitors

(University of Aberdeen) for statistical consultancy, Mary Whitmore for training in group facilitation skills, and the Scottish National Neonatal Screening Laboratory, Yorkhill, Glasgow for providing breastfeeding data at 5-7 days gathered on Guthrie cards.
Contributors: See bmj.com.
Funding: The Chief Scientist's Office of the Scottish Government Health Directorate provided funding through a research grant (CZH/4/156). PH was funded through a primary care research career award, and the Health Economics Research Unit, University of Aberdeen received core funding. The funding body had no role in the study design, data collection, analysis, or writing of reports; it received interim progress reports at six and 18 months and a final report after three years.
Competing interests: None declared.
Ethical approval: The Metropolitan Multi-centre Research Ethics Committee approved the study.
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Ip S, Cheung M, Raman G, Chew P, Magula N, DeVine D, et al. *Breastfeeding and maternal and infant health outcomes in developed countries*. Rockville, MD: Agency for Healthcare Research and Quality, 2007. (Evidence report/technology assessment report No 153.)
2 Horta BL, Bahl R, Martines JC, Victora CG. *Evidence of the long-term effects of breastfeeding*. Geneva: WHO, 2007.
3 Bolling K, Grant K, Hamlyn B, Thornton A. *Infant feeding survey 2005*. London: NHS Information Centre, 2007.
4 Dyson L, McCormick F, Renfrew MJ. Interventions for promoting the initiation of breastfeeding. *Cochrane Database Syst Rev* 2005;(2):CDCD001688.
5 Britton C, McCormick FM, Renfrew MJ, Wade A, King SE. Support for breastfeeding mothers. *Cochrane Database Syst Rev* 2007;(1):CD001141.
6 Perera R, Heneghan C, Yudkin P. Graphical method for depicting randomised trials of complex interventions. *BMJ* 2007;334:127-9.
7 Hoddinott P, Lee AJ, Pill R. Effectiveness of a breastfeeding peer coaching intervention in rural Scotland. *Birth* 2006;33:27-36.
8 Hoddinott P, Chalmers M, Pill R. One-to-one or group based peer support for breastfeeding? Women's perceptions of a breastfeeding peer coaching intervention. *Birth* 2006;33:139-46.
9 Hoddinott P, Pill R, Chalmers M. Health professionals, implementation and outcomes: reflections on a complex intervention to improve breastfeeding rates in primary care. *Fam Pract* 2007;24:84-91.
10 McGuire M, Dagge-Bell F, Purton P, Thompson M. Shaping maternity services in Scotland. *British Journal of Midwifery* 2004;12:674-8.
11 Waterman H, Tillen D, Dickson R, de Koning K. Action research: a systematic review and guidance for assessment. *Health Technol Assess* 2001;5(23):iii-157.
12 Kramer MS, Chalmers B, Hodnett ED, Sevkovskaya Z, Dzikovich I, Shapiro S, et al. Promotion of Breastfeeding Intervention Trial (PROBIT): a randomized trial in the Republic of Belarus. *JAMA* 2001;285:413-20.
13 Bartington S, Griffiths LJ, Tate AR, Dezateux C, the Millennium Cohort Study Child Health Group. Are breastfeeding rates higher among mothers delivering in baby friendly accredited maternity units in the UK? *Int J Epidemiol* 2006;35:1178-86.
14 Broadfoot M, Britten J, Tappin DM, MacKenzie JM. The baby friendly hospital initiative and breast feeding rates in Scotland. *Arch Dis Child Fetal Neonatal Ed* 2005;90:F114-6.

Accepted: 30 September 2008

Antenatal peer support workers and initiation of breast feeding: cluster randomised controlled trial

Christine MacArthur,[1] Kate Jolly,[1] Lucy Ingram,[1] Nick Freemantle,[1] Cindy-Lee Dennis,[2] Ros Hamburger,[3] Julia Brown,[3] Jackie Chambers,[3] Khalid Khan[4]

EDITORIAL by Cattaneo
RESEARCH p 388

[1]School of Health and Population Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT
[2]University of Toronto, Canada
[3]Heart of Birmingham Teaching Primary Care Trust, Birmingham
[4]School of Clinical and Experimental Medicine, University of Birmingham, and Birmingham Women's Hospital
Correspondence to: K Jolly
c.b.jolly@bham.ac.uk

Cite this as: *BMJ* 2009;338:b131
doi:10.1136/bmj.b131

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;338:b131

ABSTRACT

Objective To assess the effectiveness of an antenatal service using community based breastfeeding peer support workers on initiation of breast feeding.
Design Cluster randomised controlled trial.
Setting Community antenatal clinics in one primary care trust in a multiethnic, deprived population.
Participants 66 antenatal clinics with 2511 pregnant women: 33 clinics including 1140 women were randomised to receive the peer support worker service and 33 clinics including 1371 women were randomised to receive standard care.
Intervention An antenatal peer support worker service planned to comprise a minimum of two contacts with women to provide advice, information, and support from approximately 24 weeks' gestation within the antenatal clinic or at home. The trained peer support workers were of similar ethnic and sociodemographic backgrounds to their clinic population.
Main outcome measure Initiation of breast feeding obtained from computerised maternity records of the hospitals where women from the primary care trust delivered.
Results The sample was multiethnic, with only 9.4% of women being white British, and 70% were in the lowest 10th for deprivation. Most of the contacts with peer support workers took place in the antenatal clinics. Data on initiation of breast feeding were obtained for 2398 of 2511 (95.5%) women (1083/1140 intervention and 1315/1371 controls). The groups did not differ for initiation of breast feeding: 69.0% (747/1083) in the intervention group and 68.1% (896/1315) in the control groups; cluster adjusted odds ratio 1.11 (95% confidence interval 0.87 to 1.43). Ethnicity, parity, and mode of delivery independently predicted initiation of breast feeding, but randomisation to the peer support worker service did not.
Conclusion A universal service for initiation of breast feeding using peer support workers provided within antenatal clinics serving a multiethnic, deprived population was ineffective in increasing initiation rates.
Trial registration Current Controlled Trials ISRCTN16126175.

INTRODUCTION

In 2005 only 77% of women in England and Wales initiated breast feeding.[1] Although this has increased from 71% since 2000,[2] there is still variation across groups, with lower rates in socioeconomically deprived populations and in some ethnic minority groups. The UK government has set a target for primary care trusts to increase initiation rates for breast feeding by 2% a year. Among other interventions to achieve this, peer support is being used. We evaluated

the effectiveness of a community based antenatal service using peer support workers on initiation of breast feeding in a multiethnic deprived population.

METHODS

The general practice antenatal clinic was the unit of randomisation. The study setting was a primary care trust in a deprived area of Birmingham, with 5500-6000 deliveries per year and 90% to women from ethnic minority groups.[3] Most of the deliveries are in three hospitals (96%), with 1% at home. We included all general practices in the primary care trust in the study. In some cases more than one practice shared an antenatal clinic: for the purposes of trial allocation we considered these practices as one cluster. Randomisation was stratified by size of clinic and by midwifery team.

The intervention was a community based antenatal breastfeeding service using peer support workers developed by the primary care trust mainly to increase its initiation rate for breast feeding. The service was in addition to usual antenatal care provided by midwives. It comprised 11 peer support workers for breast feeding who were recruited to be similar to the women in their clinics on the basis of ethnicity and language and to have breast fed for several months. The peer support worker was to make initial contact in the clinic followed by a minimum of two contacts, one at 24-28 weeks' gestation and the other around 36 weeks' gestation. At least one contact was to be at home. The purpose of the consultations was to provide advice and information on the benefits of breast feeding and to support women with cultural barriers or concerns. The peer support workers kept a log of women who reached 24-28 weeks' gestation, noting those who refused support and why. For those women who had a support session the peer support worker recorded any history of infant feeding and plans for feeding before giving advice, when and where each session took place, and issues covered. Women in the control clusters received usual information and advice from midwives on breast feeding.

The primary outcome was initiation of breast feeding, defined by the infant receiving breast milk at delivery or by hospital discharge, according to hospital records. Data were obtained for women delivering in the three main hospitals during 1 February to 31 July 2007. From hospital records we obtained information on general practice indentifying code, date of delivery, age, parity, mode of delivery, ethnic group, and Townsend deprivation score.

Statistical analysis

We did statistical analyses according to the intention to treat principle. To account for extra binomial variability in both the point estimate of the effect of treatment and the confidence intervals we treated clusters (antenatal clinics) as random effects.[4] For the analysis of the primary outcome we prespecified in our statistical analysis plan a non-linear mixed model with a logit link and binomial error, including a random effect with a Gaussian error structure. In the principal model we included only the intervention group as a fixed effect and the cluster as a random effect. Missing data were not imputed. In further prespecified exploratory analyses we examined the potential impact of the midwifery team (which covered more than one practice) by adding the team delivering care as a further fixed effect. The effect of parity, ethnicity, age, deprivation score, mode of delivery, and hospital on initiation of breast feeding was also examined. We did not adjust for multiple testing. We used multiple imputation to examine the potential effects of missing data. Analyses were done in SAS version 9.1.

RESULTS

Of 66 general practice clusters (antenatal clinics), 33 were randomly allocated to the peer support service and 33 to standard antenatal care (see bmj.com). One small intervention practice closed after randomisation but before intervention. During the six months of the study 2511 women delivered in the three hospitals, 1140 (45.4%) received antenatal care in the 32 intervention practices and 1371 (54.6%) in the 33 control practices. Data on initiation of breast feeding were available for 2398 women (95.5%); 1083 (95.0%) in the intervention group and 1315 (96.0%) in the control group.

Initiation rates did not differ between intervention (69.0%) and control (68.1%) groups: cluster adjusted odds ratio 1.11 (95% confidence interval 0.87 to 1.43), P=0.40, interpractice correlation coefficient 0.07 (table). Multiple imputation techniques provided a similar result to the analysis using complete data: cluster adjusted odds ratio 1.10 (0.86 to 1.42), P=0.44.

Initiation of breast feeding varied according to several sociodemographic and delivery characteristics (see bmj.com). Initiation was lower in Heartlands Hospital, younger and older women, those who had a

Breastfeeding status in women allocated to peer support for breast feeding or to standard antenatal care by a midwife

	Peer support group		Control group		Total	
Breastfeeding status	No (%)	% of total	No (%)	% of total	No (%)	% of total
Initiated	747 (69.0)	65.5	896 (68.1)	65.4	1643 (68.5)	65.4
Not initiated	336 (31.0)	29.5	419 (31.9)	30.6	755 (31.5)	30.1
Total	1083 (100)	—	1315 (100)	—	2398 (100)	—
Not known	57	5	56	4	113	5.5
Overall total	1140	100	1371	100	2511	100

Caesarean section, and multiparous women. The lowest initiation was among white British women and the highest among African-Caribbean women. Substantial variation was found among Asian ethnic groups, with the lowest initiation among Bangladeshi women and the highest among women of Indian (subcontinent) origin. No difference was found for deprivation score, but 70% of the sample was in the lowest 10th. Multivariable analysis with adjustment for cluster showed that being from an ethnic minority group compared with being white British, and being primiparous were independently associated with an increased likelihood of initiating breast feeding. Randomisation to the peer support worker service was not (see bmj.com).

Records of a contact were available for 912 women (80.0% of deliveries during study period), and 846 (74.2%) had a support session. Of the women contacted, 64 (7%) refused a support session because they had decided to bottle feed (n=21) or breast feed (n=43). The mean duration of the first support session was 13.1 (SD 10.2) minutes, and 799 (94.4%) took place in the clinic, with only 11 (1.3%) at home. Of the 846 women who accepted a first support session, 351 (41.5%) had second session and 25 (3.0%) a third. The first session took place at a mean of 28 (SD 6.5) weeks' gestation and the second at 34.5 (SD 3.6) weeks.

DISCUSSION

This large cluster randomised controlled trial showed no effect on initiation of breast feeding of a universal community based antenatal breastfeeding peer support service in a primary care trust with a high proportion of women from ethnic minority groups and a deprived population. Peer support was chosen by the primary care trust as the option most likely to increase initiation of breast feeding among such women, as suggested by evidence into practice briefing by the UK health service.[5]

The lack of effect shown in this trial is consistent with the findings of a randomised controlled trial in one general practice in Scotland,[6] which aimed to increase the initiation and continuation of breast feeding. This report was published after the start of our trial and too recently to be included in systematic reviews. Antenatal peer support comprised one home visit, with further visits if requested. The trial included 235 unselected women, with group allocation stratified for experience of breast feeding. Initiation rates were similar—54.5% in the peer support group and 53.1% in the control group.

Other randomised controlled trials of interventions incorporating antenatal peer support have included only women considering breast feeding, with postnatal peer support to increase continuation or exclusivity as their primary purpose. A UK trial, where selection for eligibility meant that initiation of breast feeding was high, found no effect of home based peer support on breastfeeding outcomes.[7] Two small trials in the US found an effect of peer support where the intervention incorporated home based antenatal peer contact and daily postpartum peer support in hospital.[8 9]

WHAT IS ALREADY KNOWN ON THIS TOPIC

Peer or lay support is effective in prolonging exclusive breast feeding

WHAT THIS STUDY ADDS

A universal, predominantly antenatal clinic based, peer support worker service for breast feeding is ineffective in increasing initiation rates

Strengths and weaknesses of the study

Our trial is larger than any other of the peer support trials we found through a systematic search. The coverage of women was high but the intensity of the peer contact may be a limitation because this was less than planned. The service was universal, with 80% of women offered support and 74% taking up the offer. Two antenatal sessions were planned but these were attended by only 42% of women. In addition one session should have been at home but this rarely took place, and many sessions were short. More contacts might have taken place than were recorded. Despite recruiting peer support workers who were ethnically and linguistically appropriate for the population, exact matches were not possible.

Another limitation of the trial could be that data on initiation of breast feeding were obtained from maternity records, which are not generally considered as error free as data specifically collected by a research team. However, this allowed a low loss to follow-up, at only 5%, and the quality of the data was similar across groups. Although the study groups did not differ in initiation rates a 10% absolute increase occurred from the rate when the primary care trust had decided to set up the new service, probably resulting from improved data quality.

Meaning of the study

The lack of effect found from the antenatal clinic based peer support worker service evaluated in this study suggests that such a service should not be adopted as standard care. If the service had included more home based contact it might have had an effect, although in the two other UK trials[6 7] peer support was entirely home based and no improvement occurred in any breastfeeding outcomes. The service might have needed to be more intensive, and in the other UK trials contact antenatally comprised only one visit for most women, fewer than in the present trial. Perhaps the amount of advice on breast feeding and support already provided routinely in antenatal clinics in the UK allows for little additional gain from other interventions. Peer support might be more effective if targeted at specific groups, such as those women not planning to breast feed, which was around 40% of participants in this study, or those for whom routine

advice on breast feeding is less accessible because of linguistic difficulties.

Contributors: See bmj.com.
Funding: Heart of Birmingham Teaching Primary Care Trust.
Competing interests: RH, JB, JC are from the Heart of Birmingham Primary Care Trust and were involved in the employment and management of the peer support workers.
Ethical approval: Not required.

1 Bolling K, Grant C, Hamlyn B, Thornton A. *Infant feeding survey 2005*. Leeds: NHS Information Centre, 2007.
2 Hamlyn B, Brooker S, Oleinikova K, Wands S. *Infant feeding 2000*. London: Stationery Office, 2002.
3 Taylor B, Newall D. *Maternity, mortality and migration: the impact of new communities*. Edgbaston, Birmingham: Heart of Birmingham Teaching NHS Primary Care Trust. 2008. www.wmlga.gov.uk/documents/MMM%20Document_FINAL%20Webversion.pdf.
4 Collet D. *Modelling binary data*. London: Chapman and Hall, 1991.
5 Dyson L, Renfrew M, McFaden A, McCormick F, Herbert G, Thomas J. *Promotion of breastfeeding initiation and duration. Evidence into practice briefing*. London: National Institute for Health and Clinical Excellence, Jul 2006.
6 Muirhead PE, Butcher G, Rankin J, Munley A. The effect of a programme of organised and supervised peer support on the initiation and duration of breastfeeding: a randomised trial. *Br J Gen Pract* 2006;56:191-7.
7 Graffy J, Taylor J, Williams A, Eldridge S. Randomised controlled trial of support from volunteer counsellors for mothers considering breast feeding. *BMJ* 2004:328;26-31.
8 Chapman DJ, Damio G, Young S, Perez-Escamilla R. Effectiveness of breastfeeding peer counseling in a low-income, predominantly Latina population. *Arch Pediatr Adolesc Med* 2004;158:897-902.
9 Anderson AK, Damio G, Young S, Chapman DJ, Perez-Escamilla R. A randomized trial assessing the efficacy of peer counseling on exclusive breastfeeding in a predominantly Latino low-income community. *Arch Pediatr Adolesc Med* 2005;159:836-41.

Accepted: 1 November 2008

Use of primary care electronic medical record database in drug efficacy research on cardiovascular outcomes: comparison of database and randomised controlled trial findings

Richard L Tannen, Mark G Weiner, Dawei Xie

University of Pennsylvania School of Medicine, 295 John Morgan Building, 36th and Hamilton Walk, Philadelphia, PA 19104, USA
Correspondence to: R L Tannen
tannen@mail.med.upenn.edu

Cite this as: *BMJ* 2009;338:b81
doi:10.1136/bmj.b81

ABSTRACT

Objectives To determine whether observational studies that use an electronic medical record database can provide valid results of therapeutic effectiveness and to develop new methods to enhance validity.
Design Data from the UK general practice research database (GPRD) were used to replicate previously performed randomised controlled trials, to the extent that was feasible aside from randomisation.
Studies Six published randomised controlled trials.
Main outcome measure Cardiovascular outcomes measured with hazard ratios calculated with standard biostatistical methods and a new analytical technique, prior event rate ratio (PERR) adjustment.
Results In nine of 17 outcome comparisons, there were no significant differences between results of randomised controlled trials and database studies analysed using standard biostatistical methods or PERR analysis. In eight comparisons, Cox adjusted hazard ratios in the database differed significantly from the results of the randomised controlled trials, suggesting unmeasured confounding. In seven of these eight, PERR adjusted hazard ratios differed significantly from Cox adjusted hazard ratios, whereas in five they didn't differ significantly, and in three were more similar to the hazard ratio from the randomised controlled trial, yielding PERR results more similar to the randomised controlled trial than Cox ($P<0.05$).
Conclusions Although observational studies using databases are subject to unmeasured confounding, our new analytical technique (PERR), applied here to cardiovascular outcomes, worked well to identify and reduce the effects of such confounding. These results suggest that electronic medical record databases can be useful to investigate therapeutic effectiveness.

INTRODUCTION

Implementation of electronic records in clinical practice will provide an opportunity for research related to medical treatments, provided this information is compiled into robust, well designed databases and analysed with appropriate methods. Incorrect analyses could have important negative effects on medical treatment and health policy.

Two potential problems could arise in the use of medical record databases to provide reliable information concerning treatment outcomes: the quality of the data contained within the database and the ability of analyses of non-experimental observational data to provide valid results.

Considerable controversy exists over whether observational studies can provide reliable information concerning effectiveness of therapeutics.[1-6] Because of their ability to balance measured and unmeasured confounders, randomised controlled trials remain the highest level of evidence. The quality of evidence from observational studies is lower because of confounding by indication and other biases. Several comparative analyses suggest that observational studies often yield results reasonably consistent with those of randomised controlled trials. Nevertheless, there are several well documented examples where the results from observational studies were misleading.[1-3 7-9]

An important limitation applicable to previous comparative analyses is that most of the observational studies did not have rigorous inclusion and exclusion criteria,

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;338:b81

exposure definitions, and outcomes identical to the randomised controlled trials so that lack of randomisation was not the only important difference.[1 2 6 9]

We examined both the potential research value of the electronic medical record database and the validity of observational studies. We used a new analytical method, prior event rate ratio (PERR) adjustment, to enhance the validity of the results.

METHODS

GPRD database

The UK GPRD database contains information from the electronic medical records of primary care practices encompassing a representative sample of 5.7% of the UK population during 1990-2000.[10] It includes the complete primary care medical record, information on all medications prescribed, and information from outside consultants. Limitations of the database include missing data on smoking, systolic blood pressure, body mass index, limited data on admissions to hospital, and no direct link to laboratory data or death certificates.

Comparison of outcome hazard ratios in randomised controlled trials and general practice research database (GPRD)

	Death	Myocardial infarction	Stroke	CABG/PTCA
Syst-Eur				
Trial	0.86 (0.67 to 1.09)	0.70 (0.44 to 1.09)	0.58 (0.40 to 0.83)	—
GPRD-Cox	1.23 (1.00 to 1.50)*	0.74 (0.52 to 1.07)	0.68 (0.51 to 0.94)	—
WHI-intact uterus				
Trial	0.98 (0.82 to 1.15)	1.11 (0.84 to 1.47)†	1.41 (1.07 to 1.85)	1.01 (0.83 to 1.22)
GPRD-Cox	0.75 (0.65 to 0.86)*	0.95 (0.78 to 1.16)	1.23 (0.99 to 1.52)	1.15 (0.79 to 1.67)
GPRD-PERR	—	1.40 (0.87 to 2.44)	2.63 (1.38 to 7.43)	0.57 (0.22 to 1.56)
GPRD-no missing‡	0.91 (0.79 to 1.05)	—	—	—
WHI-hysterectomy				
Trial	1.01 (0.88 to 1.22)	0.89 (0.70 to 1.12)	1.39 (1.10 to 1.77)	0.93 (0.78 to 1.10)
GPRD-Cox	0.68 (0.57 to 0.81)*	0.50 (0.38 to 0.67)*	0.95 (0.74 to 1.23)*	0.59 (0.36 to 0.95)
GPRD-PERR	—	1.28 (0.69 to 2.56)§	3.06 (1.39 to 10.31)§	1.22 (0.67 to 2.42)
GPRD-no missing‡	0.82 (0.66 to 1.02)	—	—	—
4S				
Trial	0.70 (0.58 to 0.85)	0.67 (0.58 to 0.77)	0.64 (0.47 to 0.88)	0.63 (0.54 to 0.74)
GPRD-Cox	0.71 (0.53 to 0.96)	0.79 (0.61 to 1.02)	0.90 (0.63 to 1.30)	2.22 (1.80 to 2.75)*
GPRD-PERR	—	0.69 (0.51 to 0.93)	NA¶	1.00 (0.75 to 1.33)*§
HOPE				
Trial	0.84 (0.75 to 0.95)	0.79 (0.70 to 0.89)	0.68 (0.56 to .84)	0.82 (0.74 to 0.92)
GPRD-Cox	0.94 (0.85 to 1.03)	1.42 (1.23 to 1.61)*	1.16 (0.99 to 1.35)*	1.67 1.34 to 2.07)*
GPRD-PERR	—	0.62 (0.53 to 0.74)*§	0.94 (0.77 to 1.14)*	0.75 (0.56 to 1.01)§
EUROPA				
Trial	0.89 (0.77 to 1.02)	0.76 (0.66 to 0.89)	0.96 (0.72 to 1.28)	0.96 (0.85 to 1.08)
GPRD-Cox	1.06 (0.95 to 1.19)	1.36 (1.16 to 1.58)*	1.04 (0.84 to 1.29)	2.20 (1.85 to 2.62)*
GPRD-PERR	—	0.84 (0.69 to 1.01)§	0.77 (0.55 to 1.07)	1.26 (.97 to 1.62)§

CABG/PTCA=coronary artery bypass grafts or percutaneous transluminal coronary angioplasty
*Significant difference (P<0.05) compared with trial.
†Trial values for myocardial infarction reflect WHI re-analysis by age, encompassing 50-70 years.
‡Subset not missing any data for BMI, systolic blood pressure, or smoking.
§Significant difference (P<0.05) compared with GPRD Cox adjusted hazard ratio.
¶PERR could not be done because stroke was study exclusion criteria.

GPRD study protocol

We used database replications of six randomised controlled trials (see bmj.com).[11-16] As far as possible the database studies used the same inclusion and exclusion criteria, study time frame, and treatment regimen as the randomised trials.[10 17-20]

We selected subjects for inclusion in the database studies from all database subjects who met the inclusion criteria and received treatment with the study treatment during a predefined recruitment interval. The unexposed cohort was selected from all patients who met the inclusion criteria but did not receive the study drug during the recruitment interval. Unexposed patients were randomly matched by age and sex to the exposed patients and were assigned a start time identical to the matched patient.

All database studies ended on a predefined date or on outcome stop points defined in the randomised controlled trial. We analysed database studies using a simulated "intention to treat" paradigm or an "as treated" analysis.

Statistical analysis

We determined Cox unadjusted and adjusted hazard ratios for all outcomes. The adjusted hazard ratios used a predetermined set of potential confounders including demographics, medications at baseline, and medical conditions.

We also analysed results with a propensity score approach, estimated using logistic regression, which used all demographics, drug use at baseline, and identified medical conditions as confounders.[10 17 20] Analysis stratified by the propensity scores balances the treated and untreated groups with respect to the observed covariates used in estimating the propensity scores.

We used a prior event rate ratio (PERR) approach to adjust the Cox hazard ratio.[17 20] This analysis requires that neither the exposed nor unexposed patients are treated with the study drug before the start of the study. It assumes that the hazard ratio of the exposed to unexposed for a specific outcome before the start of the study reflects the combined effect of all measured and unmeasured confounders independent of treatment.

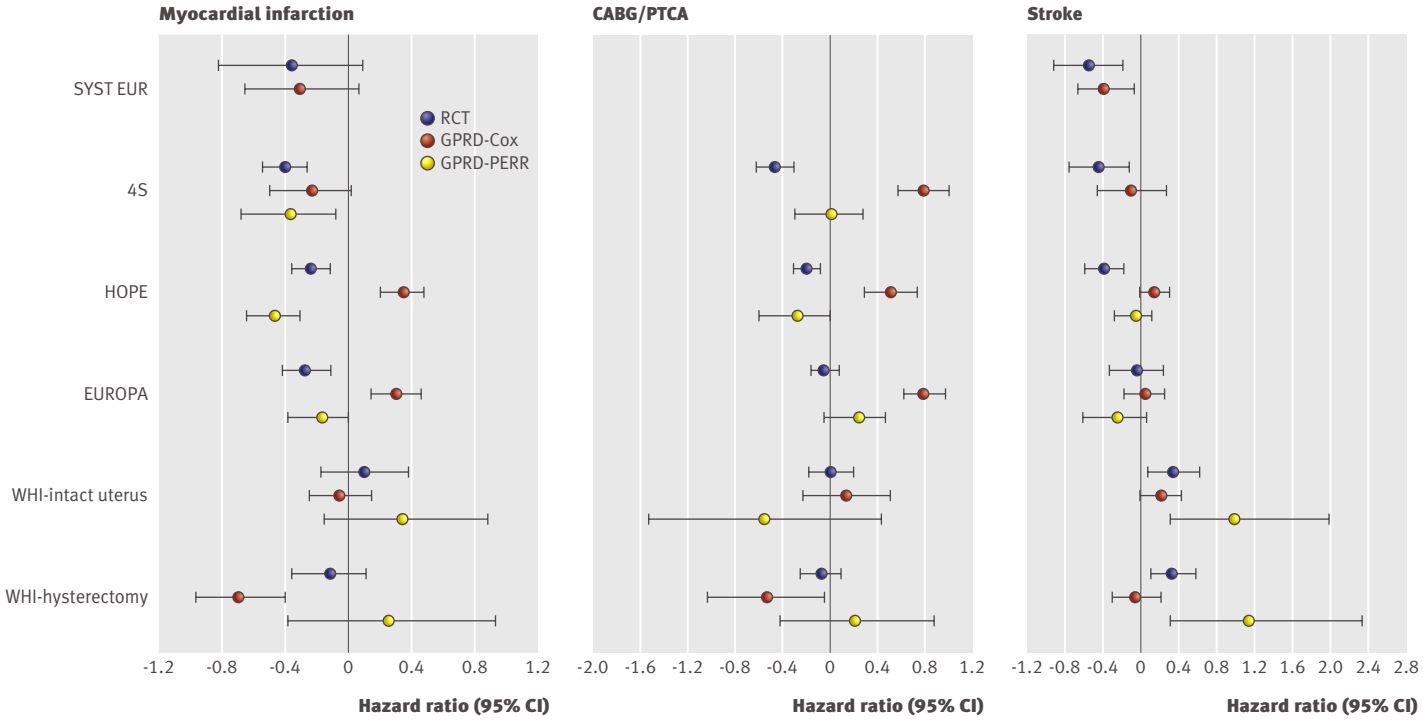
To apply the PERR adjustment method, we divided the unadjusted hazard ratio of exposed versus unexposed groups during the study by the unadjusted hazard ratio of exposed versus unexposed "before" the study. The average time of the previous period for all the outcomes assessed averaged 3.52 years (range 2.8-3.9 years).

We compared differences between the hazard ratio from the randomised trial and the database.[10]

RESULTS

Comparability

The size of the unexposed group in the database study was always larger than the placebo group of the randomised controlled trials. The exposed group in the database study was smaller than the treated cohort in half of the randomised controlled trials. The database was inadequate to replicate several randomised controlled trials because of an insufficient number of exposed patients.



Comparisons between hazard ratios from randomised controlled trials (RCT) and adjusted hazard ratios for respective database studies. Data plotted as natural logarithms, so 0 on x axis indicates no difference between exposed and unexposed cohort. Database adjusted hazard ratios shown with both Cox and prior event rate ratio (PERR) adjustment analysis. Results are shown for myocardial infarction, stroke, and coronary revascularisation (CABG/PTCA). GPRD=general practice research database

The database cohort typically differed from the respective trials in their baseline demographic characteristics, existing comorbidities, and use of cardiovascular drugs.[10 17-20] The database treatment protocol precisely replicated the trial in only one study. All but one of the database studies exhibited differences in the baseline characteristics of the exposed and unexposed groups.

Comparison of outcomes

We focused on randomised controlled trials with primary cardiovascular outcomes. We report on death, myocardial infarction, stroke, and coronary revascularisation. The table and figure show cardiovascular outcomes and statistical comparisons for the six database studies and trials. Propensity score analyses did not differ meaningfully from the analysis with Cox adjusted hazard ratios.

We compared the database studies to the six randomised controlled trials. In nine of 17 comparisons of cardiovascular outcomes there was no significant difference between the Cox adjusted hazard ratios from the database and the hazard ratios from the randomised controlled trials (table). In none of these nine comparisons did the PERR analysis differ significantly from either the trial hazard ratios or the database Cox adjusted hazard ratios.

In eight of the 17 comparisons the database Cox adjusted hazard ratios differed significantly from the trial hazard ratios, suggesting the presence of unmeasured confounding. In seven of these eight instances the PERR adjusted hazard ratios differed significantly from the Cox adjusted hazard ratios, and either did not differ significantly (five outcomes) or were more similar (two outcomes) to the trial hazard ratio. In the other outcome the PERR hazard ratio was more similar to the trial but did not differ significantly from the Cox adjusted hazard ratio. A Wilcoxon signed rank test showed that when the Cox adjusted hazard ratio differed significantly from the trial hazard ratio (n=8), the PERR adjusted hazard ratio was significantly ($P<0.05$) more similar to the trial hazard ratio than the Cox adjusted hazard ratio.

In the aggregate, when the outcome results from the database studies analysed by conventional statistical methods are confirmed or corrected by the PERR method, they largely are comparable with the results from the respective randomised controlled trials.

DISCUSSION

Our results suggest that observational studies using electronic medical record databases might produce valid results concerning the efficacy of cardiovascular drug treatments.

When analysed with conventional biostatistical analyses, the database outcome results did not differ significantly from those in the randomised controlled trial in nine of the 17 comparisons. In no instance did the PERR analysis differ significantly from the randomised controlled trial, when there was no difference between the conventional analyses and the

WHAT IS ALREADY KNOWN ON THIS TOPIC

Two major potential problems could impede the capability of an electronic medical record database to provide reliable information on drug efficacy: the quality of the data in the database and the ability of analyses of observational—that is, non-experimental—data to provide valid results

The quality of evidence from observational studies is less than from randomised controlled trials because of confounding by indication and other biases related to the effects of unmeasured covariates

WHAT THIS STUDY ADDS

Although observational studies are subject to unmeasured confounding, a new analytical technique, prior event rate ratio (PERR) adjustment, can identify and reduce unmeasured confounding

Data from properly constructed electronic medical record databases, when analysed with standard statistical methods along with the PERR method, can reveal important insights into the efficacy of medical treatment

trial. When the database outcomes analysed with conventional biostatistical techniques differed significantly from the trial, the PERR analysis results were either not significantly different from or much more similar to the trial results.

The instances where the database results analysed by conventional biostatistical methods differed importantly from the results in the trial presumably reflect unmeasured confounding by indication in the database studies. Thus our findings support concerns that the validity of observational studies must always be viewed with circumspection. The studies reported here suggest that the PERR technique can identify and largely correct for the effects of unmeasured confounding.

PERR analytical technique

The underlying hypothesis of the PERR analytical technique is that a comparison between the event rate for a specific outcome in a cohort's exposed and unexposed patients before entry into the study should reflect the effect of all confounders on that specific outcome independent of the effect of treatment. This assumption holds only when neither the exposed nor unexposed patients have been treated with the study drug before the start of the study.

When there are no unmeasured confounders, reflected by similar results of the database Cox adjusted hazard ratio and the randomised controlled trial hazard ratio, the PERR adjusted results should be similar to the Cox adjusted hazard ratio. In our analyses the PERR adjustment seemed to function in this fashion.

When there are unmeasured confounders the results of the PERR adjusted hazard ratio and the Cox adjusted hazard ratio should differ. Our results support this prediction. The PERR adjustment yielded a result more consistent with the findings in the trial, and in all but one instance, differed significantly from the Cox adjusted hazard ratio.

Further investigation is necessary to fully validate the PERR technique. More extensive statistical simulation studies would determine its limitations and applications. The technique is outcome specific, and it cannot be applied to death. Studies using databases would be supplementary to evidence from randomised controlled trials. One example might be to generalise the results of randomised controlled trials.[10 21 22]

The PERR technique should be viewed currently as applicable only to analysis of studies which include similar inclusion and exclusion criteria for the exposed and unexposed and a defined study start, recruitment interval, and end time. The random matching technique might be critical to assure that bias does not exist in the start time for unexposed patients.

Another potential shortcoming of our studies is the inability to exactly replicate all aspects of the randomised controlled trial. There is also the possibility of inaccuracy of information in the database. The similar results between the database studies and trials, however, suggest these were not major problems.

Our current view is that the PERR analysis should not be performed in isolation. We would recommend its use along with conventional biostatistical analyses. When the conventional and PERR analyses are similar, "unmeasured confounding" would seem unlikely; whereas when they differ "unmeasured confounding" would seem likely. When unmeasured confounding seems to be present, the PERR analysis seems to yield a more valid result, but additional evaluation is required to ascertain the veracity of this suggestion.

We gratefully acknowledge the assistance of Xingmei Wang, who assisted with the biostatistical analyses, and James Lewis and Stephen Kimmel for their insightful review of this manuscript.
Contributors: See bmj.com.
Funding: This work was supported by the National Institutes of Health research grant RO1-HL 073911.
Competing interests: None declared.
Ethical approval: Not required.
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C. Choosing between randomised and non-randomised studies: a systematic review. *Health Technol Assess* 1998;2:1-124.
2 MacLehose RR, Reeves BC, Harvey IM, Sheldon TA, Russell IT, Black AMS. A systematic review of comparisons of effect-sizes derived from randomized and non-randomised studies. *Health Technol Assess* 2000;4:1-154.
3 Ioannidis JPA, Haidich AB, Pappa M, Pantazis N, Kokori SI, Tektonidou MG, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA* 2001;286:821-30.
4 Kunz R, Oxman AD. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ* 1998;317:1185-90.
5 Popcock SJ, Elbourne DR. Randomized trials or observational tribulations? *N Engl J Med* 2000;342:1907-9.
6 Deeks JJ, Dinnes J, D'Amico RA, Sowden AJ, Sakarovitch C, Song F, Petticrew M, Altman DG. Evaluating non-randomised intervention studies. *Health Technol Assess* 2003;7:1-173.
7 Benson K, Hartz AJ. A comparison of observational studies and randomized controlled trials. *N Engl J Med* 2000;342:1878-86.
8 McKee M, Britton A, Black N, McPherson K, Sanderson C, Bain C. Interpreting the evidence choosing between randomised and non-randomised studies. *BMJ* 1999;319:312-5.
9 Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 2000;342:1887-92.
10 Tannen RL, Weiner MG, Marcus SM. Simulation of the Syst-Eur randomized control trial using a primary care electronic medical record was feasible. *J Clin Epidemiol* 2006;59:254-64.
11 Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in health postmenopausal women. Principal results from the Women's Health Initiative randomized controlled trial. *JAMA* 2002;288:321-33.
12 Staessen JA, Fagard R, Thijs L, Celis H, Arabidze GG, Birkenhager WH, et al. Randomised double-blind comparison of placebo and active

treatment for older patients with isolated systolic hypertension. *Lancet* 1997;350:757-64.
13 Women's Health Initiative Steering Committee. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy. The Women's Health Initiative randomized controlled trial. *JAMA* 2004;291:1701-12.
14 Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin survival study (4S). *Lancet* 1994;344:1383-9.
15 Heart Outcomes Prevention Evaluation Study Investigators. Effects of an angiotensin-converting-enzyme inhibitor, ramipril, on cardiovascular events in high risk patients. *N Engl J Med* 2000;342:145-53.
16 European trial on reduction of cardiac events with Perindopril in stable coronary artery disease investigators. Efficacy of perindopril in reduction of cardiovascular events among patients with stable coronary artery disease: randomized double-blind, placebo-controlled, multicentre trial (the EUROPA study). *Lancet* 2003;362:782-8.
17 Tannen RL, Weiner MG, Xie D, Barnhart K. A simulation using data from a primary care practice database closely replicated the women's health initiative trial. *J Clin Epidemiol* 2007;60:686-95.
18 Tannen RL, Weiner MG, Xie D, Barnhart K. Estrogen affects post-menopausal women differently than estrogen plus progestin therapy. *Hum Reprod* 2007;22:1769-77.
19 Weiner MG, Xie D, Tannen RL. Replication of the Scandinavian Simvastatin Survival Study using a primary care medical record database prompted exploration of a new method to address unmeasured confounding. *Pharmacoepidemiol Drug Saf* 2008;17:661-70.
20 Tannen RL, Weiner MG, Xie D. Replicated studies of two randomized trials of angiotensin converting enzyme inhibitors: further empiric validation of the "prior event rate ratio" to adjust for unmeasured confounding by indication. *Pharmacoepidemiol Drug Saf* 2008;17:671-85.
21 Weiner MG, Barnhart K, Xie D, Tannen RL. Hormone replacement therapy and coronary heart disease in young women. *Menopause* 2008;15:86-93.
22 Tannen RL, Weiner MG, Xie D, Barnhart K, Perspectives on hormone replacement therapy: the Women's Health Initiative and new observational studies sampling the overall population. *Fertil Steril* 2008;90:258-64.

Accepted: 19 October 2008

Preserving professional credibility: grounded theory study of medical trainees' requests for clinical support

Tara J T Kennedy,[1,2,3] Glenn Regehr,[4,5,1] G Ross Baker,[4] Lorelei Lingard[1,6,7,8]

[1]Wilson Centre for Research in Education, Toronto, ON, Canada
[2]Bloorview Kids Rehab, Toronto, ON, Canada
[3]Stan Cassidy Centre for Rehabilitation, Fredericton, NB, Canada
[4]Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[5]Toronto General Research Institute, University Health Network, Toronto, ON, Canada
[6]SickKids Learning Institute, Toronto, ON, Canada
[7]Department of Paediatrics, University of Toronto, Toronto, ON, Canada
[8]University Health Network, Toronto, ON, Canada

Correspondence to: Tara Kennedy
tara.kennedy@utoronto.ca

Cite this as: *BMJ* 2009;338:b128
doi:10.1136/bmj.b128

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;338:b128

ABSTRACT

Objective To develop a conceptual framework of the influences on medical trainees' decisions regarding requests for clinical support from a supervisor.

Design Phase 1: members of teaching teams in internal and emergency medicine were observed during regular clinical activities (216 hours) and subsequently completed brief interviews. Phase 2: 36 in-depth interviews were conducted using videotaped vignettes to probe tacit influences on decisions to request support. Data collection and analysis used grounded theory methods.

Setting Three teaching hospitals in an urban setting in Canada.

Participants 124 members of teaching teams on general internal medicine wards and in the emergency department, comprising 31 attending physicians, 57 junior and senior residents, 28 medical students, and eight nurses. Purposeful sampling to saturation was conducted.

Results Trainees' decisions about whether or not to seek clinical support were influenced by three issues: the clinical question (clinical importance, scope of practice), supervisor factors (availability, approachability), and trainee factors (skill, desire for independence, evaluation). Trainees perceived that requesting frequent/inappropriate support threatened their credibility and used rhetorical strategies to preserve credibility. These strategies included building a case for the importance of requests, saving requests for opportune moments, making a plan before requesting support, and targeting requests to specific team members.

Conclusions Trainees consider not only clinical implications but also professional credibility when requesting support from clinical supervisors. Exposing the complexity of this process provides the opportunity to make changes to training programmes to promote timely supervision and provides a framework for further exploration of the impact of clinical training on quality of care of patients.

INTRODUCTION

Quality of care in clinical training depends, to a large part, on communication between trainees and supervisors in the form of requests for clinical help when required. This works much of the time, but when it doesn't, the results can be devastating.

Ethnographic studies of medical education and research on case presentations have suggested that communication between trainees and their supervisors is not a straightforward transmission of clinical information but rather is complicated by issues like evaluation, learning agendas, and professional socialisation.[1-4]

We developed a conceptual framework of the influences on trainees' decisions to request clinical help from a supervisor through exploration of teaching team practices.[5]

METHODS

We designed the study using grounded theory. Participants were 124 members of clinical teaching teams in general internal medicine and emergency medicine, including attending physicians (n=31), junior and senior residents (n=57), third or fourth year medical students (n=28), and nurses (n=8). We used purposeful sampling[6] to ensure inclusion of participants of both sexes and of different levels of experience. Sampling continued until saturation of the data was reached.

The project took place in three academic health sciences centres associated with an urban Canadian medical school. The amount of direct contact between attending physicians and trainees was determined by the individual attending physician.[7] In Canada, attending physicians provide evaluation and feedback for all trainees under their supervision.

Phases of study

The study had two phases. The principal investigator and a research assistant collected all data; neither had any affiliation with the study sites or previous knowledge of the participants.

Phase 1 involved non-participant observation of 12 teaching teams (seven in general internal medicine and five in emergency medicine, 88 team members observed). Each team was observed six times over one month (total of 216 hours). Detailed structured field notes were kept, which evolved as the study progressed to reflect emergent analytical concepts.[6] As a source of triangulating data, 65 members of teaching teams also completed a brief interview near the end of the month of observations. The interviews were used to explore the authenticity of the observational data and to probe the intentions and rationales behind observed behaviours. The interviews were audiotaped and anonymised.

We analysed the field notes, reflective notes, and interview transcripts for emergent themes using grounded theory methods. Data collection and analysis proceeded simultaneously in an iterative fashion. Two researchers recursively read the dataset to develop a preliminary coding structure,[6] and the research team confirmed this coding structure. One coder applied the final coding structure to the complete dataset.

Phase 2 was designed to refine and expand the emerging conceptual framework, through in-depth interviews with video prompts. We developed a series of 10 videotaped vignettes, each crafted to present a dilemma, taken from the observational data, that was relevant to decisions about supervision. For example, one vignette portrayed a junior resident deliberating about whether or not she should call her attending physician before giving a patient heparin in the middle of the night.

Phase 2 participants included 19 attending physicians, 13 residents, and four medical students. We report here mainly on phase 1 data and the trainee interviews from phase 2.

RESULTS

Trainees' decisions about whether or not to seek clinical support were influenced by the nature of the clinical situation and also by factors related to supervisors and to the trainees themselves. See tables on bmj.com.

The more urgent the clinical situation, or the more important the implications of the clinical decision to be made, the more likely trainees were to seek support from their supervisors. Trainees also considered whether or not the clinical question fell within the scope of practice that was expected of them at their level of training. A senior resident explained: "If I want [attending physician] to watch me suture, then he probably would laugh at me because that is expected of me to know that" (senior resident 1, emergency medicine). Trainees also had an impression of the types of clinical situations that always warranted contacting their supervisors. For example, one medical student relayed a situation where there was a question about a patient's discharge status: "I contacted [attending physician] because . . . obviously I cannot make the call of whether or not [patient's name] can stay in the hospital" (medical student 8, general internal medicine).

Trainees also considered issues related to their supervisors. They considered their supervisor's proximity and availability. When supervisors were perceived to be busy, or were not physically present and therefore potentially busy with something else, trainees thought their clinical question had to be sufficiently important to warrant a call for help. Trainees also considered the approachability of a supervisor, or the ease with which they perceived that they could ask a particular supervisor for support. As a senior resident explained: "You get a vibe from your staff very quickly… And some staff you get the impression that if you call them in the middle of the night it's going to be a huge deal and they'll be talking in the morning and be sort of like 'I can't believe him. He called in the middle of the night' . . ." (senior resident 10, general internal medicine).

Finally, trainees' requests for help were influenced by factors related to the trainees themselves. For example, trainees described how their desire for independence in clinical skills affected their decisions about seeking help from supervisors.

Trainees also discussed how their concern about evaluation and assessment could affect decisions about asking for help. A senior resident said: "Evaluations. They figure into it… And evaluations are all subjective right? So you piss off the guy by waking him up and he's going to give you a bad evaluation. If that matters to you, you won't wake him up" (senior resident 6, emergency medicine).

Trainees also realise that they might, in some cases, lack the expertise required to recognise when they cannot cope and need support from their supervisors.

We saw that trainees struggle with (and sometimes, particularly in the middle of the night, agonise over) the decision about whether or not to call their supervisor to ask for help. What makes these decisions so difficult? The issue of credibility emerged from our data as central to this question. The fact that trainees use rhetorical strategies (persuasive arguments)[7,8] to preserve their credibility when making requests for help supports this theory. Trainees used four main rhetorical strategies to preserve their credibility when requesting clinical support from a supervisor.

Building a case—Trainees often emphasise details that support the urgency or the importance of their clinical situation as a type of justification for making the request (see bmj.com).

Saving questions—Trainees saved questions and asked them at times convenient to their supervisor. Disturbing a supervisor for an "insignificant" question was perceived as a threat to credibility, while asking the same

question of a supervisor who was close by and available did not have the same implications. See bmj.com.

Making a plan—Trainees preserved credibility by making plans to check with a supervisor, rather than asking an open ended question about what to do.

Targeted questions—Trainees often targeted questions to less "powerful" members of the team, such as junior residents or allied health professionals, to avoid exposing a lack of clinical knowledge or skill to a supervisor. See bmj.com.

DISCUSSION

During clinical training one of the main safety mechanisms is the trainees' obligation to ask for clinical help from a supervisor when faced with problems that exceed their clinical knowledge and skills. Trainees consider the preservation of their professional credibility along with the clinical implications of their situation when deciding about whether and how to ask for help. Training programmes and clinical educators should not take for granted timely requests for appropriate clinical support.

The fact that trainees are concerned about their professional credibility during communication with their supervisors does not mean they are unconcerned about patients' wellbeing. As one junior resident said: "you realise that if you make the wrong decision you can cause serious harm to someone. So you get over looking stupid really quickly and just ask for help" (junior resident 4, general internal medicine). Also, the desire to develop independence in clinical reasoning is an appropriate educational pursuit.

Implications for medical education practice

We have provided a framework for several interventions that might improve transparency in supervisor-trainee communication about supervision. For example, trainees could be explicitly taught rhetorical strategies such as "planning before asking" and "targeting questions" to increase their comfort with requesting support, or an explicit "scope of practice" could define the types of situations that would always warrant contacting a supervisor.

Our results are also important for clinical supervisors to understand and manage the factors involved in trainees' decisions about requesting support. A simple example is the issue of supervisor availability. Increasing the frequency of meetings between supervisors and trainees might promote timely discussion about trainees' clinical concerns.

Our study shows that asking for help from a supervisor involves a complex decision making process. This leads to a question of key importance to medical educators: how and why has a medical training culture evolved in which asking for help can be so difficult? Further exploration of this sociocultural phenomenon will be required to understand and influence the multifactorial contributors to a medical training culture in which asking for help is difficult.

We took two important measures to minimise the impact of observer effect on the results[9]; firstly, consistent researchers and prolonged observation periods allowed acclimatisation to the observation, and, secondly, participants were not aware of the specific focus on supervision until the observations were completed. Although data were collected on medical services in one city, the reported themes were common across the clinical settings and institutions, supporting the transferability of the analysis.[10]

In conclusion, the exposure of the complexity of medical trainees' decisions about asking for support provides the opportunity to make changes to training programmes that might promote timely discussion of trainees' clinical concerns and also provides a framework for necessary further exploration of the impact of clinical training practices on quality of care.

WHAT IS ALREADY KNOWN ON THIS TOPIC

The quality of clinical care provided to patients by medical trainees depends on the assumption that trainees can and will request clinical support from their supervisors when required

Communication between medical trainees and their supervisors is complicated by issues such as evaluations and learning agendas

WHAT THIS STUDY ADDS

Interview and observational data from clinical teaching teams showed that trainees' decisions about requesting support from clinical supervisors are complex

Trainees consider not only clinical implications but also personal credibility when making requests for clinical support

Exposure of factors affecting medical trainees' requests for clinical support provides the opportunity to make changes to clinical training programmes to promote timely clinical oversight by supervisors

We thank Douglas Buller of the Wilson Centre, University of Toronto, for his technical expertise in the production of the videos, and Lisa Kitchen of the University of Ontario Institute of Technology for her invaluable work in data collection and study coordination.

Contributors: See bmj.com.

Funding: This study was supported by an operating grant from the Canadian Institutes of Health Research (CIHR), which had no involvement in the research process. LL is supported by a CIHR New Investigators Award.

Competing interests: None declared.

Ethical approval: This study was approved by the institutional review board and informed consent was given by all participants.

1 Lingard L, Garwood K, Schryer CF, Spafford MM. A certain art of uncertainty: case presentation and the development of professional identity. *Soc Sci Med* 2003;56:603-16.
2 Sommers PS, Muller JH, Saba GW, Draisin JA, Shore WB. Reflections-on-action: medical students' accounts of their implicit beliefs and strategies in the context of one-to-one clinical teaching. *Acad Med* 1994;69:S84-6.
3 Konner M. *Becoming a doctor*. New York: Viking Penguin, 1987.
4 Sinclair S. *Making doctors*. Oxford: Oxford International Publishers, 1997.
5 Kennedy TJ, Regehr G, Baker GR, Lingard LA. Progressive independence in clinical training: a tradition worth defending? *Acad Med* 2005;80:S106-11.
6 Glaser B, Strauss A. *The discovery of grounded theory: strategies for qualitative research*. Chicago: Aldine, 1967.
7 Burke K. *Language as symbolic action: essays on life, literature, and method*. Berkeley, CA: University of California Press, 1966.
8 Lingard L, Haber RJ. Teaching and learning communication in medicine: a rhetorical approach. *Acad Med* 1999;74:507-10.
9 Holden JD. Hawthorne effects and research into professional practice. *J Eval Clin Pract* 2001;7:65-70.
10 Lincoln Y, Guba EG. *Naturalistic inquiry*. Thousand Oaks, CA: Sage, 1985.

Accepted: 7 October 2008