The *BMJ* (impact factor 12.8) 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

RESEARCH

Effect of a multimodal high intensity exercise intervention in cancer patients undergoing chemotherapy: randomised controlled trial

Lis Adamsen,[1 2] Morten Quist,[2] Christina Andersen,[2] Tom Møller,[2] Jørn Herrstedt,[3] Dorte Kronborg,[4] Marie T Baadsgaard,[2] Kirsten Vistisen,[5] Julie Midtgaard,[2] Birgitte Christiansen,[5] Maria Stage,[2] Morten T Kronborg,[6] Mikael Rørth[7]

[1]Institute of Public Health, University of Copenhagen, Copenhagen, Denmark
[2]University Hospitals Centre for Nursing and Care Research, Copenhagen University Hospital, DK-2100 Copenhagen, Denmark
[3]Department of Oncology, Odense University Hospital, Odense, Denmark
[4]Centre for Statistics, Copenhagen Business School, Frederiksberg, Denmark
[5]Department of Oncology, Herlev University Hospital, Herlev, Denmark
[6]Department of Mathematical Sciences, University of Copenhagen
[7]Department of Oncology, Copenhagen University Hospital

Correspondence to: L Adamsen, University Hospitals Centre for Nursing and Care Research, Copenhagen University Hospital, DK-2100 Copenhagen, Denmark la@ucsf.dk

Cite this as: *BMJ* 2009;339:b3410
doi: 10.1136/bmj.b3410

ABSTRACT

Objective To assess the effect of a multimodal group exercise intervention, as an adjunct to conventional care, on fatigue, physical capacity, general wellbeing, physical activity, and quality of life in patients with cancer who were undergoing adjuvant chemotherapy or treatment for advanced disease.

Design Randomised controlled trial.

Setting Two university hospitals in Copenhagen, Denmark.

Participants 269 patients with cancer; 73 men, 196 women, mean age 47 years (range 20-65) representing 21 diagnoses. Main exclusion criteria were brain or bone metastases. 235 patients completed follow-up.

Intervention Supervised exercise comprising high intensity cardiovascular and resistance training, relaxation and body awareness training, and massage, nine hours weekly for six weeks in addition to conventional care, compared with conventional care.

Main outcome measures European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30), Medical Outcomes Study Short Form (MOS SF-36), Leisure Time Physical Activity Questionnaire, muscular strength (one repetition maximum), maximum oxygen consumption (Vo_2max).

Statistical methods The general linear model was used for continuous outcome while analysis of associates between categorical outcomes was performed as analysis of marginal homogeneity in contingency tables.

Results Adjusted for baseline score, disease, and demographic covariates, the intervention group showed an estimated improvement at six weeks for the primary outcome, fatigue, of −6.6 points (95% confidence interval −12.3 to −0.9, P=0.02; effect size=0.33, 0.04 to 0.61). Significant effects were seen on vitality (effect size 0.55, 95% CI 0.27 to 0.82), physical functioning (0.37, 0.09 to 0.65), role physical (0.37, 0.10 to 0.64), role emotional (0.32, 0.05 to 0.59), and mental health (0.28, 0.02 to 0.56) scores. Improvement was noted in physical capacity: estimated mean difference between groups for maximum oxygen consumption was 0.16 l/min (95% CI 0.1 to 0.2, P<0.0001) and for muscular strength (leg press) was 29.7 kg (23.4 to 34.9, P<0.0001). No significant effect was seen on global health status/quality of life.

Conclusion A supervised multimodal exercise intervention including high and low intensity components was feasible and could safely be used in patients with various cancers who were receiving adjuvant chemotherapy or treatment for advanced disease. The intervention reduced fatigue and improved vitality, aerobic capacity, muscular strength, physical and functional activity, and emotional wellbeing, but not quality of life.

Trial registration Current Controlled trials ISRCTN05322922.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Supervised cardiovascular exercise training of moderate intensity has shown benefits in patients with breast cancer during or after adjuvant chemotherapy

Supervised group exercise interventions including patients with different cancer diagnoses and with advanced disease undergoing chemotherapy have not been evaluated in randomised clinical trials

WHAT THIS PAPER ADDS

High intensity exercise can be undertaken safely by such patients and is associated with physiological, functional, and emotional benefits

Multimodal high and low intensity exercise interventions can be offered by interdisciplinary teams of clinicians

Innovative approaches to address men's needs for physical rehabilitation in this context are required

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;339:b3410

INTRODUCTION

Fatigue is among the most frequent and burdensome side effects of chemotherapy and results in impaired or diminished physical activity.[1] Prevention and treatment of fatigue in patients undergoing chemotherapy are complicated; treatment with drugs alone is rarely adequate. Exercise training has been introduced to improve physical capacity and quality of life and to reduce fatigue. Few intervention studies have included patients who were undergoing chemotherapy and the evidence is modest.[2-4]

Typically studies investigated the effects of a single activity of moderate intensity. Additional studies are needed to provide evidence whether patients with different cancer diagnoses, stages of disease, and symptoms can benefit from combined resistance and cardiovascular training when undergoing chemotherapy.

A wide range of low intensity interventions comprising psychosocial activities are recommended as adjuvants to pharmacological therapies, to relieve nausea, pain, and fatigue and to increase the patient's perception of self control.[5 6] We developed and tested a six week multimodal group intervention consisting of high and low intensity components for men and women with cancer who were undergoing adjuvant chemotherapy or treatment for advanced disease.

METHODS

Recruitment and assignment

Participants were recruited from Rigshospitalet and Herlev Hospital, Copenhagen, from March 2004 to March 2007. Participants were eligible if they had a diagnosis of cancer, had received at least one cycle of chemotherapy, and were aged 18-65 years. Exclusion criteria were brain or bone metastases.

Baseline measures were obtained from participants, who were randomly allocated to an intervention or a control group. The patients were stratified by sex, diagnosis category (breast, bowel, other oncological malignancies, haematological malignancies), and disease status (no evidence of disease or evidence of disease). Data analysts were blinded to patient allocation.

Intervention: exercise programme

Participants assigned to the intervention group received standard medical care while participating in a group based multimodal high and low intensity exercise intervention.

The intervention consisted of three 90 min high intensity physical training sessions followed by 30 min relaxation training, plus one 90 min session of body awareness training followed by 30 min of relaxation training each week. The intervention was undertaken over a six week period for nine hours per week, equal to 43 metabolic equivalent of task (MET) hours per week.[7] The high intensity physical training sessions comprised three components: 30 min warm-up exercises, 45 min resistance training, and 15 min cardiovascular training.

Low intensity physical training comprised three psychosocial components: relaxation, body awareness and restorative training, and massage.

Control group

Participants assigned to the control group received conventional medical care and were allowed to increase physical activity freely.

Primary and secondary outcome measures

The primary outcome, fatigue, was assessed using the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30; 30 items).[8] The EORTC QLQ-C30 comprises five functional scales, nine symptom scales or items, and a global health status/quality of life scale. Secondary outcomes included all other scales on EORTC QLQ-C30, and general wellbeing was further assessed with the Medical Outcomes Study Short Form (MOS SF-36),[9] which contains eight scales measuring general health concepts and two summary scales: physical component scale and mental component scale. Leisure time physical activity level was explored by questionnaire.

Other secondary outcomes included muscular strength and aerobic capacity (Vo_2max).

Statistical analysis

We performed a regression analysis using differences in outcome between baseline and six weeks in all outcome measures (EORTC QLQ-C30, MOS SF-36, Vo_2max, and muscular strength) as the dependent variable. The variable intervention/control was fixed and the following covariates were tested: sex, age, cohabitation, educational level, baseline outcome score, relative change in β haemoglobin, Vo_2max, one repetition maximum knee extension, and the five disease related covariates—diagnosis, evidence of disease, relapse of disease, and chemotherapy cycles before and during the study period. All analyses used the intention to treat principle.

RESULTS

Participants

During the study period 1956 patients with cancer aged 18-65 years were referred to chemotherapy. Two hundred and sixty nine patients met the inclusion criteria and agreed to participate. For the primary outcome (fatigue) we obtained post-intervention data from 118 participants in the intervention group (87.4%) and 117 in the control group (87.3%). The intervention group adherence rate was 70.8% (17 of 24 training days, range 3-24) of their supervised exercise sessions.

The control and intervention groups were matched at baseline for demographic and medical characteristics. Participants were on average 47 years old (range 20-65 years), and 73% were female. The study included patients with 21 different cancer diagnoses. Forty eight percent had evidence of disease and 52% no evidence of disease.

The participants received a total of 59 different chemotherapy regimens during the study period. The intervention and control groups had received a mean of 2.5 and 2.6 cycles of chemotherapy, respectively, before the study period, and received a mean of 1.9 and 1.8 cycles, respectively, during the six week study period.

MAIN OUTCOMES

Changes in patient rated outcomes

We found a significant effect in favour of the intervention group for the primary outcome, fatigue. The fatigue score was reduced in the intervention group by an estimated mean difference of −6.6 points (95% CI −12.3 to −0.9) compared with the control group (P=0.02, effect size=0.33, 95% CI 0.04 to 0.61) (table 1).

We noted significant effects of the intervention on seven of 10 subscales on general wellbeing (MOS SF-36): physical functioning (effect size 0.37, 95% CI 0.09 to 0.65), role physical (0.37, 0.10 to 0.64), vitality (0.55, 0.27 to 0.82) role emotional (0.32 0.05 to 0.59),

Table 1 | Health related Quality of Life (EORTC QLQ-C30) outcome variables and estimated differences

Outcome variable	Mean (SD)		Test (reference: control)	
	Baseline	6 weeks	Estimated mean difference (95% CI)	P*
Maximum number				
Control	134	117	NA	NA
Intervention	135	118	NA	NA
Global health status/quality of life				
Control	60.2 (22.4)	63.3 (22.4)		
Intervention	63.8 (21.1)	67.2 (20.3)	2.2 (−2.7 to 7.1)	0.4
Physical functioning				
Control	84.0 (15.7)	86.4 (14.5)		
Intervention	84.7 (14.5)	89.0 (12.4)	2.4 (−0.4 to 5.1)	0.09
Role functioning				
Control	65.6 (28.5)	68.9 (26.5)		
Intervention	68.7 (28.4)	74.8 (26.3)	4.6 (−1.7 to 10.9)	0.2
Emotional functioning				
Control	75.7 (19.3)	80.6 (17.8)		
Intervention	77.6 (12.2)	81.3 (17.2)	−0.3 (−4.0 to 3.4)	0.9
Cognitive functioning				
Control	81.9 (17.8)	81.3 (19.8)		
Intervention	84.4 (17.1)	83.8 (16.7)	1.7 (−2.6 to 6.0)	0.4
Social functioning				
Control	78.6 (21.2)	79.4 (20.8)		
Intervention	83.3 (20.6)	82.6 (20.5)	4.5 (−1.4 to 10.3)	0.1
Fatigue				
Control	43.0 (23.9)	41.0 (22.7)		
Intervention	39.7 (25.8)	34.6 (24.3)	−6.6 (−12.3 to −0.9)	0.02
Nausea and vomiting				
Control	17.3 (21.7)	13.7 (18.4)		
Intervention	16.0 (22.9)	13.4 (17.3)	−0.6 (−5.2 to 3.9)	0.8
Pain				
Control	15.4 (22.7)	16.8 (20.6)		
Intervention	17.4 (22.4)	14.6 (17.1)	−2.9 (−7.6 to 1.7)	0.2
Dyspnoea				
Control	18.2 (25.3)	18.2 (24.4)		
Intervention	15.3 (21.6)	14.3 (21.5)	−2.9 (−8.6 to 2.7)	0.3
Insomnia				
Control	32.7 (31.5)	26.4 (27.5)		
Intervention	22.7 (26.9)	18.4 (23.9)	−3.5 (−9.5 to 2.6)	0.3
Appetite loss				
Control	16.5 (26.6)	14.6 (22.1)		
Intervention	18.4 (27.6)	12.8 (22.3)	−3.4 (−8.9 to 2.0)	0.2
Constipation				
Control	22.2 (29.5)	18.4 (25.3)		
Intervention	16.7 (24.0)	20.8 (39.4)	4.5 (−4.0 to 13.1)	0.3
Diarrhoea				
Control	14.0 (23.5)	13.0 (21.9)		
Intervention	14.8 (28.0)	17.0 (26.1)	4.1 (−1.7 to 10.0)	0.2
Financial difficulties				
Control	14.3 (27.7)	13.0 (25.9)		
Intervention	11.1 (21.5)	10.8 (19.3)	−0.3 (−5.0 to 4.4)	0.9

*Adjusted general linear model.

mental health 0.28, 0.02 to 0.56), physical component scale (0.35 0.06 to 0.63), and mental component scale (0.41, 0.14 to 0.69) (table 2). Significant interaction was seen for role physical, with greater improvement in patients with no evidence of disease than in those with evidence of disease.

Changes in objectively measured physiological outcomes

Vo_2max increased more in the intervention group than in the control group (see bmj.com); the mean improvement in Vo_2max was 10.7% (SD 0.5) compared with no change in the control group. Significant improvements in muscular strength were seen in favour of the intervention group. The average improvement in muscular strength was 29.6% (SD 36.4) for the intervention group.

DISCUSSION

Principal findings

Men and women with a broad range of ages, cancer diagnoses, disease statuses, and chemotherapy regimens participated in this study. The multimodal intervention of high intensity exercise, relaxation, body awareness training, and massage for patients undergoing chemotherapy showed broad effects. We noted significant effects on fatigue, vitality, physical functioning, role functioning, role emotional, mental health, physical component scale, and mental component scale, and physical capacity, while global health status/quality of life and symptom scales did not show improvements.

The primary outcome, fatigue, was the most frequently reported symptom; 65% of the study population reported a fatigue level greater than that of the general population at baseline[10] and 29% reported severe fatigue.

The effect size of the improvement in fatigue (0.33) suggests a small to medium clinically important change.[11-13] Our result differs from findings of a meta-analysis that indicated that the magnitude of effect of exercise on cancer related fatigue might be too small to be clinically meaningful.[2] A Cochrane meta-analysis found the association between exercise and fatigue to be insignificant and inconclusive owing to lack of studies.[4]

Severe fatigue results from extreme muscular de-conditioning caused by both the disease and treatment but can also be triggered by a sedentary lifestyle.[14] However, only 18% of the study population had a sedentary lifestyle at baseline, which may indicate that their fatigue burden was primarily due to the disease or to the chemotherapy.

Consistent with recently published studies and meta analyses on exercise interventions,[2-4 15 16] we found no significant improvements in global health status/quality of life. Two recent trials in breast cancer patients found that neither aerobic exercise nor resistance training significantly improved cancer specific quality of life or general quality of life.[13 16]

The failure of the intervention to significantly improve global health status/quality of life indicates that this type of short term intervention was not able to overcome the complexity of patients' overall negatively affected situation.

The intervention showed no significant effect on seven of eight somatic symptom scales in the EORTC QLQ-C30 questionnaire. This finding might be due to the fact that many side effects induced by chemotherapy can be prevented or treated by supportive care drugs leading to a floor effect on these measures.[17]

Significant effects of the intervention were recorded for seven of ten subscales for general wellbeing, with small to medium effect sizes for six of the scales. Particularly, vitality showed an effect size that was greater

Table 2 | General wellbeing (MOS SF-36) outcome variables and estimated differences

Outcome variable	Mean (SD)		Test (reference: control)	
	Baseline	6 weeks	Estimated mean difference (95% CI)	P*
Maximum number				
Control	134	117	NA	NA
Intervention	135	118	NA	NA
Physical functioning				
Control	83.6 (14.8)	84.3 (16.2)		
Intervention	84.3 (13.7)	88.2 (13.2)	4.4 (1.1 to 7.7)	0.01
Role physical				
Control	27.1 (35.7)	31.8 (37.6)		
Intervention	30.5 (35.2)	46.1 (40.2)	12.4 (3.4 to 21.5)	0.007
Bodily pain				
Control	74.0 (24.6)	75.7 (22.7)		
Intervention	70.9 (27.6)	77.6 (20.0)	3.4 (−1.2 to 8.0)	0.2
General health perceptions				
Control	61.1 (22.9)	65.5 (22.4)		
Intervention	64.9 (18.4)	68.7 (19.7)	0.6 (−3.7 to 4.9)	0.8
Vitality				
Control	55.8 (21.1)	55.6 (21.6)		
Intervention	57.8 (20.2)	65.5 (18.1)	8.8 (4.4 to 13.1)	<0.0001
Social functioning				
Control	75.4 (21.8)	76.5 (22.0)		
Intervention	77.0 (21.1)	79.7 (22.2)	2.4 (−2.8 to 7.6)	0.4
Role emotional				
Control	58.6 (41.2)	58.7 (41.9)		
Intervention	56.1 (39.0)	69.6 (40.1)	12.0 (1.9 to 22.0)	0.02
Mental health				
Control	72.0 (16.7)	74.2 (16.1)		
Intervention	74.0 (16.3)	78.6 (15.0)	3.3 (0.2 to 6.4)	0.04
Physical component scale				
Control	44.3 (8.3)	45.1 (8.5)		
Intervention	44.2 (8.4)	47.4 (6.7)	1.9 (0.3 to 3.4)	0.02
Mental component scale				
Control	46.9 (10.2)	47.3 (10.0)		
Intervention	46.5 (9.7)	50.5 (9.4)	3.2 (1.1 to 5.4)	0.004

*Adjusted general linear model.

than medium. Changes in fatigue were strongly affected by an increase in vitality. This finding may suggest that the multimodal intervention generates vitality and thereby reduces fatigue.

Patients in the intervention group scored high on the physical functioning scale at baseline, and after the six week intervention, their score was similar to that of the general Danish population.[18] This finding confirms possibly the patient group's predisposition for doing physical activity and that the intervention group was prepared to partake in demanding activities without health related constraints. With respect to the role emotional, mental health scales, and mental component scale, the patients in the intervention group showed significant improvement but had lower scores than their age equivalents in the general population.

Objectively measured physiological outcomes also showed significant improvement in aerobic capacity and muscular strength. Studies in healthy adults and in people with cardiac and renal illnesses and with diabetes have shown that combined resistance and cardiovascular training programmes can have a range of beneficial effects such as increased physical function, aerobic capacity, and reduced muscular fatigue.[19-23]

On days with high intensity exercise training, the sessions closed with low intensity relaxation training. The aim was to assist the patients to recognise and test their own physical reactions, such as dizziness, overexertion, and cold sweat. These findings correspond to single intervention studies regarding patients' perception of self control.[24]

Strengths and weaknesses

The strengths of the trial include: supervised and structured exercise, combined high and low intensity components, use of validated objective physiological measurements, validated questionnaires, intention to treat analyses, and limited drop-out rate of 12.7%. All participants were undergoing chemotherapy during the study period. Limitations include an adherence rate of 70.8% and a 53% recruitment rate, which are comparable with other exercise interventions including cancer patients with lesser disease burden.[15 16]

It was not possible to perform valid comparisons of the effect between the control and the intervention groups 3 months post intervention.

Self selection of participants in our study resulted in a sample of cancer patients who were overtly motivated to engage in group based physical activity. Our findings suggest that population heterogeneity does not preclude use of this type of intervention. However, generalisation may be limited by the willingness of patients to allocate the necessary time to physical activity.

Contributors: LA, CA, MQ, TM, JM, and MR contributed to conception and trial design. LA, JM, MR, and JH wrote the protocol. JH, KV, BC, CA, and TM were responsible for recruitment of patients to the trial. MQ, CA, and TM conducted the daily training sessions. CA, MQ, TM, and BC collected the data. JH, MR, and KV provided clinical expertise. DK was study statistician. MTB, MS and MTK contributed to the statistical analysis and interpretation of data. LA and MR drafted the first version of the manuscript. LA obtained the funding. JH, DK, KV, MQ, TM, CA, BC, and JM critically reviewed, revised, and supplemented the manuscript. All authors approved the final version. LA is the guarantor.

Funding: This research was supported by grants from The Lundbeck Foundation, The Novo Nordic Foundation, The Egmont Foundation, The Danish Cancer Society, The Svend Andersen Foundation, The Aase and Ejnar Danielsen Foundation, The Beckett Foundation, The Wedell-Wedellsborg Foundation, The Hede Nielsen Family Foundation, The Gangsted Foundation, Copenhagen University Hospital. The authors' work was independent from the funders.

Competing interests: None declared.

Ethical approval: Ethical approval was obtained from the Scientific Committees of the Copenhagen and Frederiksberg Municipalities (j.no. 01273/00) and the Danish Data Protection Agency (j.no. 2000-41-0-149).

1 Smets EM, Garssen B, Schuster-Uitterhoeve AL, de Haes JC. Fatigue in cancer patients. *Br J Cancer* 1993;68:220-4.
2 Schmitz KH, Holtzman J, Courneya KS, Mâsse LC, Duval S, Kane R. Controlled physical activity trials in cancer survivors: a systematic review and meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2005;14:1588-95.
3 McNeely ML, Campbell KL, Rowe BH, Klassen TP, Mackey JR, Courneya KS. Effects of exercise on breast cancer patients and survivors: a systematic review and meta-analysis. *CMAJ* 2006;175:34-41.
4 Markes M, Brockow T, Resch KL. Exercise for women receiving adjuvant therapy for breast cancer. *Cochrane Database Syst Rev* 2006;4:CD005001.
5 Kangas M, Bovbjerg DH, Montgomery GH. Cancer-related fatigue: a systematic and meta-analytic review of non-pharmacological therapies for cancer patients. *Psychol Bull* 2008;134:700-41.
6 Kwekkeboom KL, Hau H, Wanta B, Bumpus M. Patients' perceptions of the effectiveness of guided imagery and progressive muscle relaxation interventions used for cancer pain. *Complement Ther Clin Pract* 2008;14:185-94.

7 Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc* 2000;32(9 Suppl):S498-504.
8 Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;85:365-76.
9 Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-83.
10 Klee M, Groenvold M, Machin D. Quality of life of Danish women: population-based norms of the EORTC QLQ-C30. *Qual Life Res* 1997;6:27-34.
11 Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale: Lawrence Erlbaum Associates, 1988.
12 Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev Camb Philos Soc* 2007;82:591-605.
13 King MT. The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. *Qual Life Res* 1996;5:555-67.
14 Lucía A, Earnest C, Pérez M. Cancer-related fatigue: can exercise physiology assist oncologists? *Lancet Oncol* 2003;4:616-25.
15 Mutrie N, Campbell AM, Whyte F, McConnachie A, Emslie C, Lee L, et al. Benefits of supervised group exercise programme for women being treated for early stage breast cancer: pragmatic randomised controlled trial. *BMJ* 2007;334:517-20.
16 Courneya KS, Segal RJ, Mackey JR, Gelmon K, Reid RD, Friedenreich CM, et al. Effects of aerobic and resistance exercise in breast cancer patients receiving adjuvant chemotherapy: a multicenter randomized controlled trial. *J Clin Oncol* 2007;25:4396-404.
17 Roila F, Hesketh PJ, Herrstedt J; Antiemetic Subcommittee of the Multinational Association of Supportive Care in Cancer. Prevention of chemotherapy- and radiotherapy-induced emesis: results of the 2004 Perugia International Antiemetic Consensus Conference. *Ann Oncol* 2006;17:20-8.
18 Bjørner JB, Damsgaard MT, Watt T, Bech P, Rasmussen NK, Kristensen TS, et al. Dansk manual til SF-36: et spørgeskema om helbredsstatus [SF-36 (Danish manual): health-related quality of life questionnaire]. Copenhagen: Danish Association of the Pharmaceutical Industry, 1997.
19 Nader GA. Concurrent strength and endurance training: from molecules to man. *Med Sci Sports Exerc* 2006;38:1965-70.
20 Gayda M, Choquet D, Ahmaidi S. Effects of exercise training modality on skeletal muscle fatigue in men with coronary heart disease. *J Electromyogr Kinesiol* 2009;19:e32-9.
21 Moinuddin I, Leehey DJ. A comparison of aerobic exercise and resistance training in patients with and without chronic kidney disease. *Adv Chronic Kidney Dis* 2008;15:83-96.
22 Thomas DE, Elliott EJ, Naughton GA. Exercise for type 2 diabetes mellitus. *Cochrane Database Syst Rev* 2006;3:CD002968.
23 Milne HM, Wallman KE, Gordon S, Courneya KS. Effects of a combined aerobic and resistance exercise program in breast cancer survivors: a randomized controlled trial. *Breast Cancer Res Treat* 2008;108:279-88.
24 Luebbert K, Dahme B, Hasenbring M. The effectiveness of relaxation training in reducing treatment-related symptoms and improving emotional adjustment in acute non-surgical cancer treatment: a meta-analytical review. *Psychooncology* 2001;10:490-502.

Accepted: 8 May 2009

Effect of "collaborative requesting" on consent rate for organ donation: randomised controlled trial (ACRE trial)

The ACRE Trial Collaborators

Correspondence to: D Young, Intensive Care Society Clinical Trials Group, Kadoorie Centre for Critical Care Research and Education, John Radcliffe Hospital, Oxford OX3 9DU
duncan.young@nda.ox.ac.uk

Cite this as: *BMJ* 2009;339:b3911
doi: 10.1136/bmj.b3911

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;339:b3911

ABSTRACT

Objective To determine whether collaborative requesting increases consent for organ donation from the relatives of patients declared dead by criteria for brain stem death.

Design Unblinded multicentre randomised controlled trial using a sequential design. Centralised 24 hour telephone randomisation based on randomised permuted blocks of 10.

Setting 79 general, neuroscience, and paediatric intensive care units in the United Kingdom.

Participants 201 relatives of patients meeting criteria for brain stem death. Relatives were blind to the intervention and to the trial; all other participants were necessarily unblinded.

Interventions Collaborative requesting for consent for organ donation by the potential donor's clinician and a donor transplant coordinator (organ procurement officer) compared with routine requesting by the clinical team alone.

Main outcome measure Proportion of relatives consenting to organ donation.

Results 101 relatives were randomised to routine requesting and 100 to collaborative requesting. All were analysed on an intention to treat basis. In the routine requesting group, 62 relatives consented to organ donation. In the collaborative requesting group, 57 relatives consented. After correction for the ethnicity, age, and sex of the potential donors the risk adjusted ratio of the odds of consent in the collaborative requesting group relative to the routine group was 0.80 (95% confidence interval 0.43 to 1.53), with a P value of 0.49 adjusted for interim analysis and trial over-running. The conversion rate (donors with consent from whom any organs were retrieved) was 92% (57/62) in the routine requesting group and 79% (45/57) in the collaborative requesting group (P=0.043). There were 140 approaches to relatives in the per protocol analysis, leading to 60.3% (44/73) consent after routine and 67.2% (45/67) after collaborative requesting (risk adjusted odds ratio of consent 1.47, 0.67 to 3.20, P=0.33).

Conclusion There is no increase in consent rates for organ donation when collaborative requesting is used in place of routine requesting by the patient's clinician.

Trial registration ISRCTN01169903.

WHAT IS ALREADY KNOWN ON THIS TOPIC

There are several modifiable factors correlated with consent rates for organ donation

Previous research has suggested that "collaborative requesting," which involves the donor transplant coordinator in the request for organs to a potential donor's family, increases consent rates for donation

WHAT THIS STUDY ADDS

Collaborative requesting can practically be undertaken in only two thirds of requests for organ donation

Collaborative requesting has no effect on the consent rate for organ donation

INTRODUCTION

The most common reason why organs for transplantation are not obtained from patients after confirmation of brain stem death on an intensive care unit in the United Kingdom is the refusal of consent by the patient's relatives. A systematic review identified 11 observational studies suggesting that using trained and experienced individuals to make requests for organ donation increased consent rates.[1] One technique to maximise the experience of requesters is "collaborative requesting," where a request for organ donation is made jointly by the patient's clinician and a donor transplant coordinator. Although widely advocated, the efficacy or effectiveness of this technique has not been rigorously tested.

METHODS

The ACRE (Assessment of Collaborative REquesting) study was designed to test the null hypothesis that there is no difference in consent rates for organ donation when relatives are approached by the clinical team and a donor transplant coordinator together (collaborative request) compared with the clinical team alone (routine request). The study was an unblinded multicentre randomised controlled trial, with a sequential design.

Participants were the relatives of patients declared dead by criteria for brain stem death or awaiting brain stem death testing who were to be approached regarding organ donation. The study took place in 79 general, neuroscience, and paediatric intensive care units. We excluded units with in house donor transplant coordinators and a collaborative requesting rate over 50% when the study started.

Our primary outcome measure was the proportion of relatives giving consent to organ donation. Secondary outcome measures included the proportion of potential donors from whom each type of solid organ was retrieved and transplanted, and the proportion from whom tissues were retrieved.

Randomisation was made at the time when the patient's clinicians and the donor transplant coordinator agreed to request organ donation from the relatives of a patient declared dead by brain stem criteria. Relatives were blind both to the intervention and to the trial itself.

Our primary end point (consent for organ donation) was determined soon after randomisation, and because of the shortage of organs for transplantation there was a requirement not to prolong the study if a clear difference in consent rates became apparent. Consequently we used a sequential design and analysed the data using a triangular test. In this design, the extent of the difference between the consent rates in the two groups is examined after the first 100 patients have been recruited, and subsequently every 50, until there is evidence that the two rates differ or that there is no difference between them. See bmj.com. The results were analysed primarily on an intention to treat basis. Interim analyses were adjusted for the age group of the patient, ethnicity, and sex, factors that might influence the consent rate for organ donation. A per protocol analysis was also undertaken on those relatives who had a collaborative request.

Consent rates for organ donation

	All (n=201)	Routine request (n=101)	Collaborative request (n=100)
Consent to organ donation (%)	119 (59)	62	57
Any solid organ retrieved (% all patients)	102 (51)	57 (56)	45 (45)
Per protocol	140	73	67
Consent to organ donation (% per protocol patients)	89 (64)	44 (60)	45 (67)
Any solid organ retrieved (% per protocol patients)	76 (54)	39 (53)	37 (55)

RESULTS

The study recruited relatives between December 2007 and October 2008. There were no differences in characteristics of donors between the study groups, and the relatives were matched, except that there were fewer parents of donors and more children of donors in the collaborative requesting group and more patients who were registered with the UK organ donor registry or whose views on organ donation were known in the routine requesting group. The characteristics of the requesters were matched. See bmj.com.

The results of the risk adjusted sequential analysis and the planned analyses at 100 and 150 patients and the final analysis at 201 patients are shown on bmj.com.

The table shows the consent rates for all the sets of relatives by study group assignment (intention to treat). There was no difference in the rates between groups (P=0.53). Of the 201 sets of relatives, the three risk adjustment factors likely to affect consent rates were available for all but six (97%) of the potential organ donors. For these data, the risk adjusted consent rates were 58% in the collaborative requesting group and 63% in the routine group. The risk adjusted ratio of the odds of consent in the collaborative requesting group relative to the routine group was 0.80 (95% confidence interval 0.43 to 1.53) with a P value of 0.49, adjusted for the interim analysis and over-running. The results of the risk adjustment, given as odds ratios, show that consent was more likely if the patient was white (8.43 for white *v* non-white, P<0.001), female (0.60 for male *v* female, P=0.12), and in the 25-34 age group (0.85, 0.29, 1.63, 0.53, and 0.51 for 0-17, 18-24, 25-34, 35-49, 50-59 *v* ≥60, overall P=0.12).

We supplemented the intention to treat analysis with a per protocol analysis. The protocol was followed exactly in only 140 (70%) of the 201 patients. The commonest reasons why the protocol was not followed were that the families had already indicated their wishes concerning organ donation before a request was made or that brain stem death

tests revealed some evidence of brain stem activity and so organ donation was not appropriate. For the per protocol analysis, the consent rates were 67% (45/67) in the collaborative requesting group and 60% (44/73) in the routine group. The risk adjusted ratio of the odds of consent was 1.47 (0.67 to 3.20, P=0.33), obtained from patients for whom age, sex, and ethnicity were known.

The results of the analyses of secondary outcome measures are on bmj.com. There was a slightly lower conversion rate (the number of donors from whom solid organs were actually retrieved as a proportion of donors in whom consent for donation had been obtained) in the collaborative requesting group compared with the routine requesting group (79% (45) *v* 92% (57), P=0.043).

DISCUSSION

Findings in the context of existing knowledge

We found no evidence for an increase in rates of consent for organ donation from relatives when collaborative requesting was used in place of routine requesting by the patient's clinician. There was weak evidence that the presence of a donor transplant coordinator at the interview was associated with a reduction in the number of organs retrieved from donors in whom consent for donation was available. The study confirmed previous UK findings that consent was more likely if the patient was white.[2]

The consent rate for all relatives approached was 59%, almost exactly the figure reported in the large, UK-wide, potential donor audit.[3] The percentage of donors with consent for organ donation from whom solid organs were actually obtained was 86% in our study, again close to the 90% reported previously in the UK. Finally, the proportion of white potential organ donors at 92% was similar to the 93% found in the audit. We believe our results are therefore generalisable to the whole UK potential donor pool.

Our results are at variance with the results of many observational studies.[1] There are several possible reasons for this. Collaborative requesting took place in only 73% of assigned cases, but in nearly all of the remaining cases collaborative requesting would have been pointless as the relatives' views were known or requesting was logistically impossible. Any modest effect of collaborative requesting, however, would have been diluted by these cases in the intention to treat analysis. The slight excess of patients in the routine requesting group whose positive view of organ donation was known might also have diluted a modest benefit from collaborative requesting.

The most likely reason for the discrepancy with published case series is simply that collaborative requesting confers little or no advantage in requests for organ donation. Why the undoubted extra experience of the donor transplant coordinators in interviewing the relatives of potential organ donors conveys no benefit is unclear. Relatives' decisions might be largely made on the basis of long held beliefs that will not be modified by any requesting technique. The clinician or potential donor's nurse has usually had prolonged contact with the relatives of potential organ donors, and the incremental benefit of a newly introduced donor transplant coordinator's experience might be modest. There is no structured training in making requests for organ donation in the UK, and the additional practical experience of the donor transplant coordinators might require some structure to make it more effective in requests for organ donation.[4] Finally, it might be simply that donor transplant coordinators, though experienced in interviewing families in nearly all matters pertaining to organ donation, actually have limited experience in making the initial request.

The lower numbers of all specific solid organs retrieved in the collaborative requesting group, resulting in the lower conversion rate, was unexpected.

Findings in the context of UK policy and practice

The results do have some implication for policy in the UK. In 2007-8 an additional 10 in-house donor transplant coordinator posts were created with the goal of increasing the consent rate for organ donation.[5] The report of the UK Department of Health's organ donation task force has also recommended the placement of "embedded" coordinators in all hospitals in the UK.[6] In the light of our results it might be more effective to focus the efforts of these additional transplant coordinators on strategies other than collaborative requesting if they are to increase consent rates.

Details of the trial steering committee, the data monitoring and ethics committee, principal investigators and donor transplant coordinators, and regional ACRE representatives are on bmj.com.

Contributors: See bmj.com.

Funding: This study was funded by NHS Blood and Transplant.

The study sponsor (Oxford University) had no role in any aspect of the design or implementation of the study. The funder's role in study design was limited to the practicalities of setting up the randomisation service and linking the study data with the National Transplant database. The study funder took no part in the interpretation of data, writing the article and the decision to submit it for publication. All the researchers (TSC) had access to the data and approved the final analysis and manuscript.

Competing interests: Dave Collett and Karen Morgan are employed by the study funder, NHS Blood and Transplant. Chris Danbury and Susan Richards serve on the donor advisory group organised by NHS Blood and Transplant.

Ethical approval: The study was approved by Oxford REC A (reference No 06/Q1604/119), with a waiver of the requirement for consent as explained in the text.

1 Simpkin AL, Robertson LC, Barber VS, Young JD. Modifiable factors influencing relatives' decision to offer organ donation: systematic review. *BMJ* 2009;338:b991.
2 Murphy C, Counter C. Potential donor audit. Summary report for the 24 month period 1st April 2007-31st March 2009: UK Blood and Transplant Authority, 2009. www.uktransplant.org.uk/ukt/statistics/potential_donor_audit/pdf/pda_summary_report_2007-2009.pdf.
3 Barber K, Falvey S, Hamilton C, Collett D, Rudge C. Potential for organ donation in the United Kingdom: audit of intensive care records. *BMJ* 2006;332:1124-7.
4 Shafer TJ. Improving relatives' consent to organ donation. *BMJ* 2009;338:b701.
5 NHS Blood and Transplant. Annual business plan 2007/08. 2007. www.nhsbt.nhs.uk/downloads/business_plan_200708.pdf.
6 Department of Health. Organs for transplants: a report from the Organ Donation Taskforce. 2008. www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_082122.

Accepted: 29 May 2009

Common mental disorder and obesity—insight from four repeat measures over 19 years: prospective Whitehall II cohort study

Mika Kivimäki,[1 2] Debbie A Lawlor,[3] Archana Singh-Manoux,[1 4] G David Batty,[5 6 7] Jane E Ferrie,[1] Martin J Shipley,[1] Hermann Nabi,[4] Séverine Sabia,[4] Michael G Marmot,[1] Markus Jokela[1 2]

EDITORIAL by Atlantis and colleagues

[1]Department of Epidemiology and Public Health, University College London, London WC1E 6BT
[2]Finnish Institute of Occupational Health, Helsinki, Finland
[3]MRC Centre for Causal Analyses in Translational Epidemiology, University of Bristol, Bristol
[4]INSERM U687-IFR69, Hôpital Paul Brousse, Villejuif Cedex, France
[5]MRC Social and Public Health Sciences Unit, University of Glasgow, Glasgow
[6]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh
[7]The George Institute for International Health, University of Sydney, Sydney, Australia
Correspondence to: M Kivimäki
m.kivimaki@ucl.ac.uk

Cite this as: *BMJ* 2009;339:b3765
doi: 10.1136/bmj.b3765

ABSTRACT

Objectives To examine potential reciprocal associations between common mental disorders and obesity, and to assess whether dose-response relations exist.

Design Prospective cohort study with four measures of common mental disorders and obesity over 19 years (Whitehall II study).

Setting Civil service departments in London.

Participants 4363 adults (28% female, mean age 44 years at baseline).

Main outcome Common mental disorder defined as general health questionnaire "caseness;" overweight and obesity based on Word Health Organization definitions.

Results In models adjusted for age, sex, and body mass index at baseline, odds ratios for obesity at the fourth screening were 1.33 (95% confidence interval 1.00 to 1.77), 1.64 (1.13 to 2.36), and 2.01 (1.21 to 3.34) for participants with common mental disorder at one, two, or three preceding screenings compared with people free from common mental disorder (P for trend<0.001). The corresponding mean differences in body mass index at the most recent screening were 0.20, 0.31, and 0.50 (P for trend<0.001). These associations remained after adjustment for baseline characteristics related to mental health and exclusion of participants who were obese at baseline. In addition, obesity predicted future risk of common mental disorder, again with evidence of a dose-response relation (P for trend=0.02, multivariable model). However, this association was lost when people with common mental disorder at baseline were excluded (P for trend=0.33).

Conclusions These findings suggest that in British adults the direction of association between common mental disorders and obesity is from common mental disorder to increased future risk of obesity. This association is cumulative such that people with chronic or repeat episodes of common mental disorder are particularly at risk of weight gain.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Whether common mental disorders predict obesity or whether obesity predicts common mental disorders is unclear

Few studies have the repeat measures necessary to understand the nature of the association

WHAT THIS STUDY ADDS

Common mental disorder predicted subsequent obesity in a cohort of 4000 men and women over a 19 year follow-up period

Weight gain and the risk of obesity increased in a dose-response fashion with the number of episodes of common mental disorder

Little evidence existed to suggest that obesity was a risk factor for subsequent common mental disorder in people with no pre-existing mental disorder

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;339:b3765

INTRODUCTION

Common mental disorders, such as depression and anxiety, have been hypothesised to increase the risk of obesity, but the evidence is inconclusive.[1] Some studies show higher rates of obesity in people with mental health problems,[2-5] whereas others report no association or a reverse effect.[4 6 7] As the development of obesity has a relatively long induction period, chronic or recurrent mental disorder is a more plausible cause than transient mental disorder. However, previous studies have typically measured common mental disorder at only one time point.

Other series of studies have explored the opposite causal direction—that is, obesity as a risk factor for future mental disorder. Again, the results from these studies have been contradictory.[6 8-13]

Several mechanisms could explain these two causal directions. Stress and common mental disorders themselves are associated with disordered eating,[14 15] and commonly used treatments for depression have known side effects that also could lead to weight gain.[16] On the other hand, obesity may generate stigma resulting in negative feelings and common mental disorder, depending on societal attitudes towards body size.

Prospective studies with repeat measurements of both common mental disorder and obesity offer a unique opportunity for in-depth study. This study takes advantage of four waves of medical screening of British government employees to determine the direction and possible dose-response nature of the association between common mental disorder and obesity.

METHODS

Study population

The target population of the British Whitehall II study was all London based office staff, aged 35-55, working in 20 government departments at study baseline in 1985-8.[17] With a response of 73%, the baseline cohort consisted of 10 308 employees (6895 men and 3413 women).

Design

Since the first medical examination (phase 1, 1985-8), follow-up screenings with a medical examination have

taken place on three occasions over a 19 year period: phase 3 (1991-3), phase 5 (1997-9), and phase 7 (2003-4). All these phases included a standardised assessment of common mental disorder and direct measurement of weight and height.

Baseline characteristics

Demographic characteristics, behavioural risk factors, and chronic health conditions were measured. These included smoking, alcohol consumption, physical activity, blood pressure, cholesterol, diabetes, and coronary heart disease.[17]

Common mental disorder

We assessed common mental disorders by using the self administered general health questionnaire,[18] a screening instrument designed for community settings. Response categories are scored as either 1 or 0 to indicate whether a symptom is present or not. We defined people with a total score of 5 or more as general health questionnaire cases and those scoring 0-4 as non-cases.[19]

Obesity and overweight

We calculated body mass index. We considered participants with a body mass index of 25-29.9 to be overweight and those with a body mass index of 30 or above to be obese.[20]

Statistical analysis

We ran two sets of analyses to examine the dose-response pattern and direction of the association between common mental disorders and obesity.

Analysis 1: common mental disorder as a risk factor for subsequent overweight and obesity—We divided participants into four groups on the basis of the number of times they were general health questionnaire cases at the first three measurement phases (phases 1, 3, and 5). We used multinomial logistic regression analysis to summarise the association between occurrence of "caseness" and body mass index category at the end of follow-up (phase 7).

Analysis 2: obesity as a risk factor for future common mental disorder—We divided participants into four groups on the basis of the number of times they were obese at the first three measurement phases. We examined the associations between occurrence of obesity and general health questionnaire caseness at the end of follow-up (phase 7) by using logistic regression analysis.

In both analyses, we estimated regression estimates firstly without adjustments and then with control for age, sex, and baseline body mass index category (analysis 1) or baseline general health questionnaire caseness (analysis 2). Further adjustments took into account all baseline characteristics.

Sensitivity analyses

To examine temporality in the associations, we excluded from the analyses people who were obese or overweight at baseline (analysis 1) and people who were general health questionnaire cases at baseline (analysis 2).

RESULTS

Sample selection and sample characteristics

The study included 10 308 participants at phase 1. Our main analysis was based on 4363 participants (3122 men and 1241 women) who had complete data on common mental disorders, weight, and height at baseline and all follow-up measurements (44.6% of the baseline cohort alive at follow-up). The mean age of the included participants was 44.0 years, and 26.7% met the criteria for general health questionnaire caseness at phase 1.

Analysis 1: common mental disorder as a risk factor for subsequent overweight and obesity

The mean follow-up from phase 1 to phase 7 was 16.5 (range 14.6-18.8) years. Of the participants included in the analysis, 2379 (54.5%) never met the definition for general health questionnaire caseness across the first three screenings, 1159 (26.6%) were cases at one screening, 585 (13.4%) were cases at two screenings, and 240 (5.5%) were cases at all three screenings.

Occurrence of general health questionnaire caseness was associated with subsequent obesity (compared with being of normal weight) in a dose-response manner: unadjusted odds ratios were 1.11 (95% confidence interval 0.90 to 1.36), 1.27 (0.98 to 1.66), and 1.56 (1.10 to 2.21) for caseness at one, two, and three screenings compared with participants who were case-free (P for trend=0.004). We found no robust association between general health questionnaire caseness at the first three screenings and subsequent overweight in the unadjusted analysis (P for trend=0.83). Table 1 shows multivariate models for these associations. The dose-response association with obesity was robust to adjustment for a wide range of potential covariates: age, sex, and baseline body mass index category (model A); plus ethnicity, marital status, socioeconomic position, smoking, alcohol intake, physical activity, blood pressure, total cholesterol, diabetes, coronary heart disease, and use of psychotropic drugs (model B).

Table 1 | Multivariate models of association between occurrence of common mental disorder (GHQ caseness) and subsequent BMI category (n=4154)

No of times GHQ case at phases 1, 3, and 5	No of participants	Odds ratio (95% CI) for overweight* *v* normal weight at phase 7 (model A†)	Odds ratio (95% CI) for obese* *v* normal weight at phase 7	
			Model A†	Model B‡
0	2259	1.00 (reference)	1.00 (reference)	1.00 (reference)
1	1111	1.19 (0.99 to 1.41)	1.33 (1.00 to 1.77)	1.31 (0.98 to 1.76)
2	555	1.32 (1.05 to 1.66)	1.64 (1.13 to 2.36)	1.63 (1.12 to 2.37)
3	229	1.00 (0.71 to 1.40)	2.01 (1.21 to 3.34)	2.01 (1.19 to 3.39)
Test for trend		P=0.07	P<0.001	P<0.001

BMI=body mass index; GHQ=general health questionnaire (measure of common mental disorder).
*1518 (36.5%) normal weight; 1938 (46.7%) overweight; 698 (16.8%) obese.
†Adjusted for age, sex, and BMI category at phase 1.
‡As model A, but additionally adjusted for ethnicity, marital status, socioeconomic position, smoking, alcohol intake, physical activity, systolic blood pressure, diastolic blood pressure, total cholesterol, diabetes, coronary heart disease, and use of psychotropic drugs at phase 1.

Table 2 | Multivariate models of association between obesity and subsequent common mental disorder (GHQ caseness; n=4154)

No of times obese at phases 1, 3, and 5	No of participants	Odds ratio (95% CI) for GHQ caseness* at phase 7	
		Model A†	Model B‡
0	3557	1.00 (reference)	1.00 (reference)
1	272	1.21 (0.89 to 1.62)	1.21 (0.89 to 1.63)
2	166	1.35 (0.93 to 1.97)	1.36 (0.93 to 1.99)
3	156	1.43 (0.98 to 2.10)	1.39 (0.94 to 2.05)
Test for trend		P=0.01	P=0.02

GHQ=general health questionnaire (measure of common mental disorder).
*834 (20.1%) were GHQ cases at phase 7.
†Adjusted for age, sex, and GHQ caseness at phase 1.
‡As model A, but additionally adjusted for ethnicity, marital status, socioeconomic position, smoking, alcohol intake, physical activity, systolic blood pressure, diastolic blood pressure, total cholesterol, diabetes, coronary heart disease, and use of psychotropic drugs at phase 1.

Analysis 2: obesity as a risk factor for future common mental disorder

Of the 4363 participants, 3727 (85.4%) were not obese at any of the three phases; 291 (6.7%) were obese once, 175 (4.0%) were obese twice, and 170 (3.9%) were obese on three occasions. The odds ratios of general health questionnaire caseness at phase 7 increased with the number of times participants were obese. Compared with non-obese people, the odds ratios in unadjusted analysis for those who were obese one, two, or three times were 1.29 (0.97 to 1.71), 1.35 (0.95 to 1.93), and 1.45 (1.02 to 2.06) (P for trend=0.004). However, as shown in multivariate models (table 2), all the odds ratios between obesity categories and subsequent general health questionnaire caseness were non-significant when adjusted for covariables, although the overall trend remained (P for trend=0.01 in model A and 0.02 in model B).

Sensitivity analyses

The dose-response relation between occurrence of caseness and subsequent obesity remained after exclusion of participants who were obese at phase 1 (n=214, P for trend<0.001) and those who were overweight (n=1356, P for trend=0.03). By contrast, the association between occurrence of obesity and subsequent general health questionnaire caseness was lost when participants defined as cases at phase 1 (n=1164) were excluded from the analysis (P for trend=0.33).

DISCUSSION

Findings from this occupational cohort of more than 4000 men and women provide evidence that common mental disorder may be a risk factor for future obesity. Additionally, we found evidence of a dose-response effect in the association of common mental disorder with obesity. Thus, the odds ratio of being obese at the fourth and final screening were two times higher for participants with common mental disorder at all three preceding screenings than for those who were disorder-free at all previous screenings. This dose-response association was independent of a wide range of mental health related characteristics at baseline that may be associated with both obesity and common mental disorder.

When we examined the association in the opposite direction (that is, obesity as a risk factor for common mental disorder), we also found a dose-response association. However, when we removed participants with common mental disorder at baseline no association of obesity with future risk of common mental disorder existed. These results suggest a temporal sequence from mental disorder to obesity in this cohort of British adults.

Strengths and weaknesses

We are not aware of any other study that has four repeat measurements of common mental disorders and direct assessment of weight and height over an extended follow-up period. However, our study uses observational data, which can never provide complete proof of causality. Even though adjustment for a wide range of covariates had little effect on our estimates of association, the possibility of residual confounding cannot be excluded. Despite a high response to the survey (range 66-88%) at the successive data collection phases, loss to follow-up accumulated over the extended time period, as is inevitable in all long term prospective studies. Furthermore, as these findings are from an occupational cohort, they may not apply to people not in paid employment.

Comparison with previous studies

Our findings are in agreement with a recent longitudinal study of elderly people that showed weight gain to be more common among people with symptoms of depression than among those who were symptom-free.[5] Our results are also consistent with those from another cohort that explored the bidirectional association between depression and type 2 diabetes, a condition associated with raised body mass index.[10] In that study, baseline depressive symptoms were associated with increased risk of type 2 diabetes at the follow-up (two years later), after removal of people with diabetes at baseline and adjustment for a wide range of potential confounding factors.

Meaning of the study

Several plausible explanations exist for the observation that persistent common mental disorder is a risk factor for obesity. Firstly, common mental disorders are associated with eating disorders, both over-consumption and under-consumption, which could influence future changes in adiposity.[14 15] Secondly, physical inactivity, a major contributing factor to obesity, is more prevalent among people with mental health problems.[21] Thirdly, commonly used drug treatments for mental disorders have known side effects that may result in weight gain (tricyclic antidepressants), weight loss (selective serotonin reuptake inhibitors), or both (short term and long term effects of selective serotonin reuptake inhibitors).[16 22] Fourthly, biological factors, such as dysregulation of the hypothalamic-pituitary-adrenocortical axis, may strengthen the link between mental disorders and obesity.[23] All these mechanisms are likely to increase the risk of obesity in a dose-response fashion.

Unanswered questions

Further longitudinal research is needed to confirm the generalisability of these findings; to identify major social, behavioural, and biological mechanisms underlying the observed associations; to identify the time lag between exposure and outcome; and to determine whether our findings are transferable to patients with specific diagnosed mental disorders. If the observed associations are causal, our findings have important implications for prevention and treatment. An increased risk of obesity should be taken into account in the treatment of persistent common mental disorders.

We thank all participating civil service departments and their welfare personnel and establishment officers; the Occupational Health and Safety Agency; the Council of Civil Service Unions; all participating civil servants in the Whitehall II study; and all members of the Whitehall II Study Team. The Whitehall II Study Team comprises research scientists, statisticians, study coordinators, nurses, data managers, administrative assistants, and data entry staff, who make the study possible.

Contributors: See bmj.com.

Funding: The Whitehall II study has been supported by grants from the Medical Research Council; British Heart Foundation; Health and Safety Executive; Department of Health; National Heart Lung and Blood Institute, US, NIH (R01HL036310-20A2); National Institute on Aging, US, NIH (R01AG034454-01 and R01AG013196); and the John D and Catherine T MacArthur Foundation Research Networks on Successful Midlife Development and Socio-economic Status and Health. MK is supported by the Academy of Finland; GDB is a Wellcome Trust research fellow; AS-M is supported by a "EURYI" award from the European Science Foundation; JEF is supported by the Medical Research Council; MJS is supported by the British Heart Foundation; and MGM is supported by an MRC research professorship. The funders did not contribute to the study design and had no role in data collection, data analysis, data interpretation, or writing of the report. The researchers were independent from the funders.

Competing interests: None declared.

Ethical approval: The University College London Medical School committee on the ethics of human research gave ethical approval for the Whitehall II study.

1 McElroy SL, Kotwal R, Malhotra S, Nelson EB, Keck PE, Nemeroff CB. Are mood disorders and obesity related? A review for the mental health professional. *J Clin Psychiatry* 2004;65:634-51, quiz 730.
2 Goodman E, Whitaker RC. A prospective study of the role of depression in the development and persistence of adolescent obesity. *Pediatrics* 2002;110:497-504.
3 Weil E, Wachterman M, McCarthy EP, Davis RB, O'Day B, Iezzoni LI, et al. Obesity among adults with disabling conditions. *JAMA* 2002;288:1265-8.
4 Stunkard AJ, Faith MS, Allison KC. Depression and obesity. *Biol Psychiatry* 2003;54:330-7.
5 Vogelzangs N, Kritchevsky SB, Beekman AT, Newman AB, Satterfield S, Simonsick EM, et al. Depressive symptoms and change in abdominal obesity in older persons. *Arch Gen Psychiatry* 2008;65:1386-93.
6 Roberts RE, Deleger S, Strawbridge WJ, Kaplan GA. Prospective association between obesity and depression: evidence from the Alameda County Study. *Int J Obes Relat Metab Disord* 2003;27:514-21.
7 Heo M, Pietrobelli A, Fontaine KR, Sirey JA, Faith MS. Depressive mood and obesity in US adults: comparison and moderation by sex, age, and race. *Int J Obes (Lond)* 2006;30:513-9.
8 Atlantis E, Baker M. Obesity effects on depression: systematic review of epidemiological studies. *Int J Obes (Lond)* 2008;32:881-91.
9 Farmer A, Korszun A, Owen MJ, Craddock N, Jones L, Jones I, et al. Medical disorders in people with recurrent depression. *Br J Psychiatry* 2008;192:351-5.
10 Golden SH, Lazo M, Carnethon M, Bertoni AG, Schreiner PJ, Roux AV, et al. Examining a bidirectional association between depressive symptoms and diabetes. *JAMA* 2008;299:2751-9.
11 Istvan J, Zavela K, Weidner G. Body weight and psychological distress in NHANES I. *Int J Obes Relat Metab Disord* 1992;16:999-1003.
12 Crisp AH, McGuiness B. Jolly fat: relation between obesity and psychoneurosis in general population. *BMJ* 1976;1:7-9.
13 Magnusson PK, Rasmussen F, Lawlor DA, Tynelius P, Gunnell D. Association of body mass index with suicide mortality: a prospective cohort study of more than one million men. *Am J Epidemiol* 2006;163:1-8.
14 Pallister E, Waller G. Anxiety in the eating disorders: understanding the overlap. *Clin Psychol Rev* 2008;28:366-86.
15 Dallman MF, Pecoraro N, Akana SF, La Fleur SE, Gomez F, Houshyar H, et al. Chronic stress and obesity: a new view of "comfort food". *Proc Natl Acad Sci U S A* 2003;100:11696-701.
16 Sussman N, Ginsberg DL, Bikoff J. Effects of nefazodone on body weight: a pooled analysis of selective serotonin reuptake inhibitor- and imipramine-controlled trials. *J Clin Psychiatry* 2001;62:256-60.
17 Marmot MG, Davey Smith G, Stansfeld S, Patel C, North F, Head J, et al. Health inequalities among British civil servants: the Whitehall II study. *Lancet* 1991;337:1387-93.
18 Goldberg DP. *Detecting psychiatric illness by questionnaire*. Oxford University Press, 1972.
19 Stansfeld SA, Marmot MG. Social class and minor psychiatric disorder in British civil servants: a validated screening survey using the general health questionnaire. *Psychol Med* 1992;22:739-49.
20 World Health Organization. *Physical status: the use and interpretation of anthropometry: report of a WHO expert committee*. WHO, 1995. (WHO Technical Report Series.)
21 Mead GE, Morley W, Campbell P, Greig CA, McMurdo M, Lawlor DA. Exercise for depression. *Cochrane Database Syst Rev* 2008(4):CD004366.
22 Demyttenaere K, Jaspers L. Review: bupropion and SSRI-induced side effects. *J Psychopharmacol* 2008;22:792-804.
23 Bornstein SR, Schuppenies A, Wong ML, Licinio J. Approaching the shared biology of obesity and depression: the stress axis as the locus of gene-environment interactions. *Mol Psychiatry* 2006;11:892-902.

Accepted: 7 July 2009

Is a simple "Thank you" too much to ask?

It is courteous and respectful to extend thanks to the people without whom the final research findings of our studies could not have been generated—that is, the patients.

I recently conducted a review of multinational randomised clinical trials and economic evaluations published over the past decade and explored the acknowledgments section of the main clinical paper looking for words of gratitude to patients. To my surprise only five (9%) of the 54 clinical studies included in the review had thanked patients for participating in the study. I also looked at randomised controlled trials published in the *BMJ* during 2009 and found that, from the 32 studies published, 13 (41%) had not thanked the patients in the manuscript.

Most of these studies included extensive lists of acknowledgments to trial investigators, editorial staff, and people who contributed to the success of the trial, but the authors from these studies seem to have overlooked their patients.

I am sure lead investigators thank patients through newsletters, information sheets, and other means, but it is the final publication that most readers study.

Of course, my estimates are not really representative of all clinical trials, but I believe patients participating in these studies deserved those encouraging words.

Oliver Rivero-Arias research officer, Health Economics Research Centre, Department of Public Health, University of Oxford, Oxford
oliver.rivero@dphpc.ox.ac.uk

I thank all patients who participate in clinical trials and are often forgotten in the acknowledgments of important manuscripts. I am grateful to Helen Campbell and Alison Gater, at the University of Oxford, and Professor Simon Eckermann, Flinders University, for their constructive and useful comments when I was preparing this note.

Cite this as: *BMJ* 2009;339:b3683

Contribution of smoking during pregnancy to inequalities in stillbirth and infant death in Scotland 1994-2003: retrospective population based study using hospital maternity records

Ron Gray,[1] Sandra R Bonellie,[2] James Chalmers,[3] Ian Greer,[4] Stephen Jarvis,[5] Jennifer J Kurinczuk,[1] Claire Williams[2]

[1]National Perinatal Epidemiology Unit, University of Oxford, Oxford OX3 7LF
[2]School of Accountancy, Economics and Statistics, Napier University, Edinburgh EH10 5DT
[3]Information Services Division (ISD), NHS National Services Scotland, Edinburgh EH12 9EB
[4]Hull York Medical School, University of York, York YO10 5DD
[5]Sir James Spence Institute of Child Health, Royal Victoria Infirmary, Newcastle upon Tyne NE1 4LP
Correspondence to: R Gray
ron.gray@npeu.ox.ac.uk

Cite this as: *BMJ* 2009;339:b3754
doi: 10.1136/bmj.b3754

ABSTRACT

Objective To quantify the contribution of smoking during pregnancy to social inequalities in stillbirth and infant death.

Design Population based retrospective cohort study.

Setting Scottish hospitals between 1994 and 2003.

Participants Records of 529 317 singleton live births and 2699 stillbirths delivered at 24-44 weeks' gestation in Scotland from 1994 to 2003.

Main outcome measures Rates of stillbirth and infant, neonatal, and post-neonatal death for each deprivation category (fifths of postcode sector Carstairs-Morris scores); contribution of smoking during pregnancy ("no," "yes," or "not known") in explaining social inequalities in these outcomes.

Results The stillbirth rate increased from 3.8 per 1000 in the least deprived group to 5.9 per 1000 in the most deprived group. For infant deaths, the rate increased from 3.2 per 1000 in the least deprived group to 5.4 per 1000 in the most deprived group. Stillbirths were 56% more likely (odds ratio 1.56, 95% confidence interval 1.38 to 1.77) and infant deaths were 72% more likely (1.72, 1.50 to 1.97) in the most deprived compared with the least deprived category. Smoking during pregnancy accounted for 38% of the inequality in stillbirths and 31% of the inequality in infant deaths.

Conclusions Both tackling smoking during pregnancy and reducing infants' exposure to tobacco smoke in the postnatal environment may help to reduce stillbirths and infant deaths overall and to reduce the socioeconomic inequalities in stillbirths and infant deaths perhaps by as much as 30-40%. However, action on smoking on its own is unlikely to be sufficient and other measures to improve the social circumstances, social support, and health of mothers and infants are needed.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Rates of both stillbirth and infant death show social gradients within developed countries

Smoking during pregnancy has been clearly linked to stillbirth and infant deaths

Quantifying the contribution that smoking during pregnancy has on the social inequalities gap in stillbirths and infant deaths is of interest

WHAT THIS STUDY ADDS

Social gradients existed in the stillbirth and infant death rates in Scotland during 1994-2003

Smoking during pregnancy accounted for 38% of the inequality in stillbirths and 31% of the inequality in infant deaths

In addition to tackling smoking during pregnancy and reducing infants' exposure to tobacco smoke, other measures are needed to reduce social inequalities in these outcomes

INTRODUCTION

Both stillbirth and infant mortality show a social gradient within developed countries.[1-6] As smoking during pregnancy has been clearly linked to stillbirths and infant deaths,[7-10] and as smoking rates during pregnancy vary markedly with socioeconomic position,[11] jointly exploring the effects of smoking and of socioeconomic position is of interest. This will answer the question of how much of an effect smoking during pregnancy has on the social inequalities gap in stillbirths and infant deaths. Few studies have examined this question directly.[12-14] We have previously shown that smoking during pregnancy may have contributed to some of the social gradients in preterm birth in Scotland.[15] Here, we examine the effects of smoking on social gradients in stillbirths and infant deaths and whether smoking during pregnancy contributed to these gradients.

METHODS

Information on all maternity admissions to all Scottish hospitals is recorded and collated on a national database. This database contains information on demographic characteristics, clinical details of care, and data on birth outcomes. This database allows the construction of a pregnancy and birth cohort study of all maternities in a general population of just over five million people containing around one million women of reproductive age and 55 000 births a year.

We extracted records of all stillbirths and live singleton births delivered at 24-44 weeks' gestation in 1994-2003 (n=541 557). We included data on 529 317 live births and 2699 stillbirths. Smoking during pregnancy was coded as "no," "yes," or "not known." Infant deaths were subdivided into death of a liveborn infant within the first 28 days of life (neonatal deaths) and death of a liveborn infant after the first 28 days of life but before 1 year of age (post-neonatal deaths). Preterm delivery was defined as delivery at less than 37 weeks' completed gestation. Socioeconomic status was categorised by using quintiles of area based deprivation scores derived from postcode sector Carstairs-Morris scores for the whole population (using 2001 census data).[16]

We determined the number of live births, stillbirths, and infant deaths in each deprivation category as well as the corresponding rates of stillbirth and neonatal, post-neonatal, and infant deaths. We calculated the odds ratios for stillbirth and for neonatal, post-neonatal,

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;339:b3754

Table 1 | Numbers and rates of stillbirths and neonatal, post-neonatal, and infant death by smoking category in Scotland 1994-2003 for 532 016 births

Smoking group	(No of births)	Rate (No) of deaths			
		Stillbirths*	Neonatal†	Post-neonatal†	Infant†
No	344 502	3.8 (1313)	2.1 (731)	1.0 (343)	3.1 (1074)
Yes	137 303	6.7 (914)	3.0 (414)	2.8 (380)	5.8 (794)
Not known	50 211	9.4 (472)	4.1 (205)	2.2 (109)	6.1 (314)
All	532 016	5.1 (2699)	2.6 (1350)	1.6 (832)	4.1 (2182)

*Rate per 1000 total births.
†Rate per 1000 live births.

Table 2 | Odds ratios (95% confidence intervals) of stillbirths and neonatal, post-neonatal, and infant deaths by deprivation category in Scotland 1994-2003, with contribution of smoking to the social inequality

Deprivation category	Stillbirths (n=2699)	Neonatal (n=1350)	Post-neonatal (n=832)	Infant (n=2182)
1 (least) (reference group):				
Unadjusted	1	1	1	1
Model A*	1	1	1	1
Model B†	1	1	1	1
2:				
Unadjusted	1.24 (1.08 to 1.42)	0.99 (0.82 to 1.20)	1.17 (0.88 to 1.55)	1.04 (0.89 to 1.22)
Model A*	1.22 (1.06 to 1.40)	0.97 (0.80 to 1.17)	1.10 (0.83 to 1.45)	1.01 (0.86 to 1.18)
Model B†	1.16 (1.01 to 1.33)	0.95 (0.78 to 1.14)	1.03 (0.78 to 1.37)	0.97 (0.83 to 1.13)
3:				
Unadjusted	1.42 (1.24 to 1.61)	1.19 (0.99 to 1.42)	1.60 (1.23 to 2.07)	1.31 (1.13 to 1.51)
Model A*	1.39 (1.22 to 1.59)	1.14 (0.95 to 1.37)	1.42 (1.09 to 1.85)	1.22 (1.05 to 1.42)
Model B†	1.27 (1.11 to 1.45)	1.09 (0.91 to 1.30)	1.28 (0.98 to 1.66)	1.14 (0.98 to 1.32)
4:				
Unadjusted	1.46 (1.28 to 1.66)	1.10 (0.92 to 1.32)	1.95 (1.52 to 2.50)	1.35 (1.17 to 1.56)
Model A*	1.42 (1.25 to 1.62)	1.05 (0.87 to 1.26)	1.65 (1.29 to 2.13)	1.23 (1.06 to 1.42)
Model B†	1.28 (1.12 to 1.46)	0.99 (0.82 to 1.19)	1.44 (1.12 to 1.86)	1.13 (0.97 to 1.31)
5 (most):				
Unadjusted	1.56 (1.38 to 1.77)	1.37 (1.16 to 1.62)	2.56 (2.02 to 3.24)	1.72 (1.50 to 1.97)
Model A*	1.52 (1.34 to 1.73)	1.28 (1.07 to 1.51)	2.05 (1.61 to 2.61)	1.51 (1.32 to 1.74)
Model B†	1.32 (1.16 to 1.50)	1.19 (1.00 to 1.41)	1.70 (1.33 to 2.17)	1.35 (1.17 to 1.55)
Tests for trend‡ (P value):				
Unadjusted	<0.001	<0.001	<0.001	<0.001
Model A*	<0.001	0.003	<0.001	<0.001
Model B†	<0.001	0.04	<0.001	<0.001
Contribution of smoking (%)§	−38.46	−32.14	−33.33	−31.17

*Adjusted for year of birth, maternal age, parity, infant sex, primary obstetric intervention.
†Adjusted for year of birth, maternal age, parity, infant sex, primary obstetric intervention, smoking during pregnancy.
‡χ^2 test of effect of fitting deprivation category as ordinal rather than categorical variable.
§Contribution of smoking to inequality between deprivation categories 5 and 1.

and infant death, comparing each deprivation category with the least deprived (reference) category. We used multiple logistic regression modelling to derive odds ratios adjusted for maternal age, parity, infant's sex, and obstetric intervention (model A) and for maternal age, parity, infant's sex, obstetric intervention, and smoking status (model B). We calculated the contribution of smoking in explaining the social inequalities in outcome between the most and the least deprived categories.[17]

RESULTS

The most deprived mothers tended to be younger and to be more likely to smoke and to give birth to preterm or low birthweight babies. Equally, the least deprived mothers were more likely to be older, non-smokers, and less likely to give birth to preterm or low birthweight babies. Women in the most deprived category were three times more likely to smoke during pregnancy than were those in the least deprived category (38% *v* 13%).

Overall, 2699 stillbirths and 2182 infant deaths occurred, corresponding to rates of 5.1 stillbirths per 1000 total births and 4.1 infant deaths per 1000 live births. However, for stillbirths, the rate increased from 3.8 per 1000 in the least deprived group to 5.9 per 1000 in the most deprived group. We found a similar pattern in the neonatal, post-neonatal, and infant death rates (see bmj.com). The rates of stillbirth were lowest in the non-smokers and highest in the smoking "not known" category (table 1). We saw a similar pattern in the infant death rate and neonatal death rate.

Comparing mothers in the most deprived category with those in the least deprived category showed that stillbirths were 56% (odds ratio 1.56, 95% confidence interval 1.38 to 1.77) more likely and that infant deaths were 72% (1.72, 1.50 to 1.97) more likely in the most deprived group (table 2). Although an increased risk of neonatal death also existed in this group (37%), the really striking finding is that post-neonatal deaths were two and half times more likely in the most deprived group (odds ratio 2.56, 2.02 to3.24).

After adjustment for the differences in age distributions, parity, infant sex, and differences in obstetric intervention, the social gradient was slightly attenuated but still marked and statistically significant. After further adjustment for smoking during pregnancy, some attenuation in the gradient again occurred but a clear and significant trend was still apparent. Comparing mothers in the most deprived category with those in the least deprived category showed that stillbirths were 32% more likely and infant deaths were 35% more likely in the most deprived group. Assessment of the contribution of smoking to the inequality between the most and least deprived groups indicated that smoking accounted for about a third of the inequality: 38% of stillbirths and 31% of infant deaths.

DISCUSSION

We did a large population based retrospective cohort study of 532 016 births in Scotland between 1994-2003, of which 2699 were stillbirths and 2182 resulted in an infant death. We found that the social inequalities in stillbirth and infant death were partly but not fully attenuated by adjustment for smoking during pregnancy, which accounted for approximately a third of the inequality in stillbirths and infant deaths between the most and least deprived groups in the population. The gradient for neonatal deaths was less marked than those for stillbirths and post-neonatal deaths.

Strengths and weaknesses

The strength of our study is that it is population based and used one of the few national databases of routinely collected information. The coverage and quality of the data have been established to be good, and information on socioeconomic position and smoking is available. The use of area deprivation indices as measures of social inequality is well supported in the

literature,[18] and may be particularly useful in pregnant women, for whom determination of socioeconomic position is known to be problematic.[19] However, an area based score does not always correspond to individually measured socioeconomic position.

Previous research has indicated that women may stop, reduce, and sometimes restart smoking at various points during pregnancy.[20] In addition, many women may be reluctant to disclose a perceived socially undesirable behaviour to their clinicians during pregnancy. As a result, recording the maternal report on smoking (yes/no) during pregnancy once at the initial booking visit will give only a snapshot. We consider that by including a non-responder category in this study we have to some extent tackled the problem of misclassification of smoking status. Our analyses show that women in this "not known" group have a similar risk profile to smokers. This suggests either that they are smokers who choose not to declare themselves as such or else that their "not known" status is a marker of a risk factor that has a similar magnitude to smoking.

The apparent effects of smoking during pregnancy on post-neonatal deaths could be explained by a direct effect on the feto-placental unit during pregnancy, influencing or programming neonatal health. However, smoking during pregnancy may also act as a "marker" for smoking in the postnatal period, which is an established risk for post-neonatal death.[21] As we did not have a measure of smoking in the postnatal period, we cannot assess the interplay of these factors.

Another limitation of the data was the lack of information on the use of alcohol and illicit drugs during pregnancy and on maternal pre-pregnancy weight. All these factors are related to both smoking and stillbirth/infant death and could be confounders.

Possible mechanisms and implications for clinicians and policymakers

Despite recent policy interest in infant deaths,[22] stillbirth accounts for a larger proportion of losses than do infant deaths, and in most cases the cause remains unclear. Reducing smoking during pregnancy may be one of the few modifiable risk factors for this outcome that might also reduce the social gradient. Although the contribution of NHS smoking cessation services to reducing smoking prevalence has been small, it has had a disproportionate effect in the most disadvantaged areas, thus potentially reducing inequalities.[23] Nevertheless, we would agree that more powerful and innovative targeted interventions are needed as well as action to strengthen tobacco control policy.[24]

Finally, even after taking smoking into account, most of the effects of social deprivation are unexplained. Therefore, action on smoking, although necessary, is not in itself sufficient; other measures to improve the social circumstances, social support, and health of mothers and infants are needed.

We thank Edmund Hey for helpful comments on an earlier draft.

Contributors: See bmj.com.

Funding: Partly funded by Chief Scientist Office project grant CZH/4/293. The funder had no part in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Competing interests: None declared.

Ethical approval: This was an analysis of routinely collected data for which Information Services Division (ISD) Scotland has overall permission. Permission for the necessary record linkage was granted by ISD's Privacy Advisory Committee.

1 Chalmers I. Short, Black, Baird, Himsworth, and social class differences in fetal and neonatal mortality rates. *BMJ* 1985;291:231-3.
2 Parsons L, Duley L, Alberman E. Socio-economic and ethnic factors in stillbirth and neonatal mortality in the NE Thames Regional Health Authority (NETRHA) 1983. *Br J Obstet Gynaecol* 1990;97:237-44.
3 Finch BK. Early origins of the gradient: the relationship between socioeconomic status and infant mortality in the United States. *Demography* 2003;40:675-99.
4 Olsen O, Madsen M. Effects of maternal education on infant mortality and stillbirths in Denmark. *Scand J Public Health* 1999;27:128-36.
5 Jørgensen T, Mortensen LH, Andersen AM. Social inequality in fetal and perinatal mortality in the Nordic countries. *Scand J Public Health* 2008;36:635-49.
6 Singh GK, Kogan MD. Persistent socioeconomic disparities in infant, neonatal, and postneonatal mortality rates in the United States, 1969-2001. *Pediatrics* 2007;119:e928-39.
7 Salihu HM, Wilson RE. Epidemiology of prenatal smoking and perinatal outcomes. *Early Hum Dev* 2007;83:713-20.
8 Högberg L, Cnattingius S. The influence of maternal smoking habits on the risk of subsequent stillbirth: is there a causal relation? *Br J Obstet Gynaecol* 2007;114:699-704.
9 Fleming P, Blair PS. Sudden infant death syndrome and parental smoking. *Early Hum Dev* 2007;83:721-5.
10 Kleinman JC, Pierre MB, Madans JH. The effects of maternal smoking on fetal and infant mortality. *Am J Epidemiol* 1988;127:274-82.
11 Schneider S, Schütz J. Who smokes during pregnancy? A systematic literature review of population-based surveys conducted in developed countries between 1997 and 2006. *Eur J Contracept Reprod Health Care* 2008;13:138-47.
12 Rush D, Cassano P. Relationship of cigarette smoking and social class to birth weight and perinatal mortality among all births in Britain, 5-11 April 1970. *J Epidemiol Community Health* 1983;37:249-55.
13 Stephansson O, Dickman PW, Johansson ALV, Cnattingius S. The influence of socio-economic status on stillbirth risk in Sweden. *Int J Epidemiol* 2001;30:1296-301.
14 Goy J, Dodds L, Rosenberg MW, King WD. Health-risk behaviours: examining social disparities in the occurrence of stillbirth. *Paediatr Perinat Epidemiol* 2008;22:314-20.
15 Gray R, Bonellie SR, Chalmers J, Greer I, Jarvis S, Williams C. Social inequalities in preterm birth in Scotland 1980-2003: findings from an area-based measure of deprivation. *Br J Obstet Gynaecol* 2008;115:82-90.
16 Carstairs V, Morris R. *Deprivation and health in Scotland*. Aberdeen University Press, 1991.
17 Singh-Manoux A, Nabi H, Shipley M, Guéguen A, Sabia S, Dugravot A, et al. The role of conventional risk factors in explaining social inequalities in coronary heart disease: the relative and absolute approaches to risk. *Epidemiology* 2008;19:599-605.
18 Galobardes B, Lynch J, Davey Smith G. Measuring socioeconomic position in health research. *Br Med Bull* 2007;81-82:21-37.
19 Fairley L, Leyland AH. Social class inequalities in perinatal outcomes: Scotland 1980-2000. *J Epidemiol Community Health* 2006;60:31-6.
20 Pickett KE, Rathouz PJ, Kasza K, Wakschlag LS, Wright R. Self-reported smoking, cotinine levels, and patterns of smoking in pregnancy. *Paediatr Perinat Epidemiol* 2005;19:368-76.
21 Moon RY, Horne RS, Hauck FR. Sudden infant death syndrome. *Lancet* 2007;370:1578-87.
22 Department of Health. *Implementation plan for reducing health inequalities in infant mortality: a good practice guide*. DH, 2007.
23 Bauld L, Judge K, Platt S. Assessing the impact of smoking cessation services on reducing health inequalities in England: observational study. *Tob Control* 2007;16:400-4.
24 Steinberg MB, Schmelzer AC, Richardson DL, Foulds J. The case for treating tobacco dependence as a chronic disease. *Ann Intern Med* 2008;148:554-6.

Accepted: 9 May 2009

pico

Cost effectiveness analysis of including boys in a human papillomavirus vaccination programme in the United States

Jane J Kim, Sue J Goldie

EDITORIAL by Castle and Scarinci

Harvard School of Public Health, Department of Health Policy and Management, Center for Health Decision Science, 718 Huntington Avenue, Boston, MA 02115, USA

Correspondence to: J J Kim
jkim@hsph.harvard.edu

Cite this as: *BMJ* 2009;339:b3884
doi: 10.1136/bmj.b3884

STUDY QUESTION Is it cost effective to include preadolescent boys (aged 12 years) in a human papillomavirus (HPV) vaccination programme for preadolescent girls in the United States?

SUMMARY ANSWER Given currently available information, our analysis suggests that if vaccine coverage and efficacy are high among preadolescent girls, then including boys in an HPV vaccination programme is unlikely to be cost effective, even under favourable conditions of vaccine protection and health benefits.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS Guidelines for HPV vaccination in the US prioritise girls aged 11 to 12 and may include boys in the near future. This study found that under conditions of high vaccine coverage and efficacy in girls, including boys in an HPV vaccination programme generally exceeds conventional thresholds of good value for money.

Main results

With 75% vaccination coverage and complete, lifelong vaccine efficacy, routine HPV vaccination of 12 year old girls was consistently less than $50 000 per quality adjusted life year (QALY) compared with screening alone. Adding 12 year old boys to the vaccination programme cost nearly $300 000 per QALY when including benefits related to cervical disease only and fell below $100 000 per QALY only under scenarios of high, lifelong vaccine efficacy against all HPV related conditions (including other non-cervical cancers, genital warts, and juvenile onset recurrent respiratory papillomatosis), or scenarios of lower vaccine efficacy with lower coverage or lower vaccine costs. Lowering vaccine efficacy in males from 90% to 75% resulted in less attractive cost effectiveness ratios that exceeded $100 000 per QALY.

Design

We conducted a cost effectiveness analysis from the societal perspective.

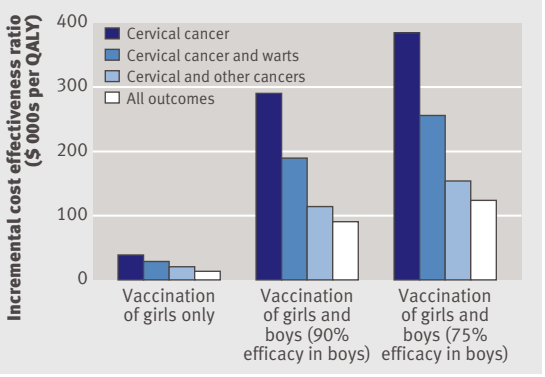
Source(s) of effectiveness

Estimates of vaccine efficacy and screening test characteristics were obtained from clinical studies and randomised controlled trials reported in the published literature.

Data sources

Model variables were based on epidemiological studies, cancer registries, population surveys, and demographic statistics from the US.

IMPACT OF DECREASED HPV VACCINE EFFICACY IN BOYS ON COST EFFECTIVENESS RATIOS



Results of sensitivity analysis

Results were stable across a range of scenarios for cervical cancer screening. When we assumed lower vaccine efficacy, waning immunity, or higher vaccine costs, vaccination of boys consistently exceeded $250 000 per QALY when reflecting benefits to cervical disease only and $100 000 per QALY when including other HPV related conditions. When lower vaccine efficacy was accompanied by lower achievable coverage or lower vaccine cost, the cost of vaccinating boys fell below $100 000 per QALY when all outcomes for both sexes were included. Costs associated with screening and treatment for cervical cancer and cross protective effects of the vaccine on cervical disease related to oncogenic HPV types other than HPV 16 and HPV 18 had minimal impact on the overall results.

Limitations

Limitations include uncertainties in the disease course of HPV related conditions and vaccine properties in the long term. We assumed an optimistic coverage rate for routine vaccination of both sexes and did not consider temporary catch-up scenarios. We did not incorporate decrements in quality of life associated with overdiagnosis of precancerous lesions attributable to screening or potential adverse events attributable to vaccination. Our analysis does not address decision making at the individual level.

Study funding/potential competing interests

The authors are supported by the National Cancer Institute (R01 CA93435), the Centers for Disease Control and Prevention, and the American Cancer Society. We have no competing interests.

This is a summary of a paper that was published on bmj.com as *BMJ* 2009;339:b3884

pico

Incidence of pregnancy after expectant, medical, or surgical management of spontaneous first trimester miscarriage: long term follow-up of miscarriage treatment (MIST) randomised controlled trial

Lindsay F P Smith,[1] Paul D Ewings,[2] Catherine Quinlan[1]

[1]East Somerset Research Consortium, Westlake Surgery, West Coker, Somerset BA22 9AH
[2]Research and Development Support Unit, Taunton and Somerset Hospital, Taunton TA1 5DA
Correspondence to: L Smith
research@esrec.nhs.uk

Cite this as: *BMJ* 2009;339:b3827
doi: 10.1136/bmj.b3827

STUDY QUESTION Does management method for early miscarriage affect long term fertility rates?

SUMMARY ANSWER Method of miscarriage management does not affect subsequent pregnancy rates with around four in five women giving birth within five years of the index miscarriage.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS After miscarriage women are increasingly offered the option of expectant or medical management as well as traditional surgical care, but the long term effects of the three methods in terms of subsequent fertility are not known. The trial shows that long term fertility is not affected by the method of management.

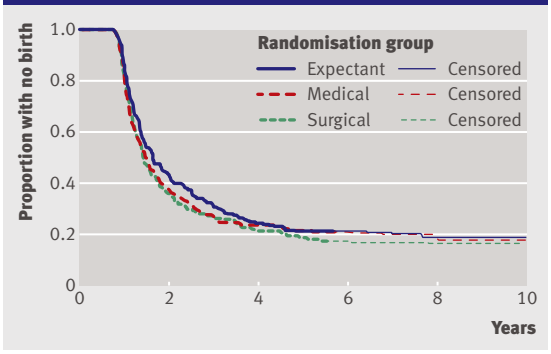
Design
Women were randomly allocated by central telephone randomisation to either surgical evacuation of retained products of conception; medical treatment with mifepristone or misoprostol, or both; or expectant management. Blinding to intervention was not possible.

Participants and setting
1200 women who had experienced an early miscarriage (<13 weeks), confirmed by scan, were recruited between 1997 and 2001 from the early pregnancy clinics of acute hospitals in the south west region of England.

Primary outcome(s)
First self reported live birth after index miscarriage.

Main results and the role of chance
Completed questionnaires were returned for 68% of the women; respondents were representative of the original trial population. Five years after the index miscarriage 177/224 (79.0%, 95% confidence interval 73.2% to 83.8%) of those who had been randomised to expectant management had given birth compared with 181/230 (78.7%, 73.0% to 83.5%) of the medical group and 192/235 (81.7%, 76.3% to 86.1%) of the surgical group (P=0.41 for log rank comparison of time to birth). Having had one or more previous pregnancies before the index miscarriage did not significantly predict time to giving birth (P=0.10). Time to subsequent birth, however, was predicted by maternal age (P<0.001) and having had a previous miscarriage (P<0.001). Five years after the index miscarriage, 378/447 (84.6%) of those women with no previous miscarriage had given birth; the corresponding figures for 1, 2, and ≥3 previous miscarriages were 122/166 (73.5%), 33/49 (67.3%), and 14/24 (58.3%), respectively.

TIME TO LIVE BIRTH AFTER MISCARRIAGE ACCORDING TO MANAGEMENT METHOD



Harms
Bleeding after the index miscarriage was higher in those randomised to expectant or medical management. Risk of unplanned admission was highest in the expectant arm, and duration of hospital stay was highest in the medical arm.

Bias, confounding, and other reasons for caution
We cannot be sure that the non-responders to the follow-up questionnaire had similar live birth rates to the responders, consequently the overall rate might in reality be different from that reported. Response rates, however, were similar for the three groups, and it is unlikely that any difference in birth rates between responders and non-responders would be sufficiently different between the three groups to negate the main finding of similar rates in the groups.

Generalisability to other populations
The study was conducted in one region of England, but there is no obvious reason why the results would not generalise to other areas and populations with a different composition—for example, concerning ethnic mix.

Study funding/potential competing interests
The study was funded by the Claire Wand Fund of the BMA.

Trial registration number
National Research Register N0467011677/N0467073587.

This is a summary of a paper that was published on bmj.com as *BMJ* 2009;339:b3827

pico

Hazardous cosleeping environments and risk factors amenable to change: case-control study of SIDS in south west England

Peter S Blair,[1] Peter Sidebotham,[2] Carol Evason-Coombe,[1] Margaret Edmonds,[1] Ellen M A Heckstall-Smith,[1] Peter Fleming[1]

EDITORIAL by Mitchell

[1]Department of Community Based Medicine, University of Bristol
[2]Health Sciences Research Institute, University of Warwick
Correspondence to: P Fleming, FSID Research Unit, St Michael's Hospital, Bristol BS2 8EG
peter.fleming@bris.ac.uk

Cite this as: *BMJ* 2009;339:b3666
doi: 10.1136/bmj.b3666

This is a summary of a paper that was published on bmj.com as *BMJ* 2009;339:b3666

STUDY QUESTION What are the circumstances in which sudden infant death syndrome (SIDS) occurs?

SUMMARY ANSWER Many of the SIDS infants had coslept in a hazardous environment. The major risk factors, regardless of markers for socioeconomic deprivation, are amenable to change: particularly never cosleeping on a sofa, and avoiding alcohol or drugs before cosleeping in a bed.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS A higher proportion of the residual deaths from SIDS now occurs among more deprived families and those who cosleep with their infant. Many of the deaths while cosleeping occurred in potentially hazardous environments, including on a sofa or shared surface with an adult who had consumed alcohol or drugs.

Participants and setting

Our study population, set in the south west of England (184 800 births), comprised 80 SIDS infants and two control groups weighted for age and time of reference sleep: 87 randomly selected controls and 82 controls at high risk of SIDS (young, socially deprived, multiparous mothers who smoked).

Design, size, and duration

This was a population based case-control study over four years. Parents were interviewed shortly after the death or after the reference sleep (within 24 hours) for the two control families.

Primary outcome(s), risks, exposures

Both distal and proximal risk factors associated with SIDS, in particular those risk factors associated with the infant sleeping environment before death or the reference sleep. The primary exposure of interest was the cosleeping environment.

Main results and the role of chance

The median age at death (66 days) was more than three weeks less than in a study in the same region a decade earlier. Of the SIDS infants, 54% died while cosleeping compared with 20% among both control groups. Much of this excess is explained by a significant multivariable interaction between cosleeping and recent parental use of alcohol or drugs (31% *v* 3% random controls) and the increased proportion of SIDS infants who had coslept on a sofa (17% *v* 1%). More SIDS infants than random controls used a pillow for the last sleep (21% *v* 3%), were swaddled (24% *v* 6%), were found prone (29% *v* 10%), were preterm (26% *v* 5%) or were in fair or poor health for the last sleep (28% *v* 6%). More mothers of SIDS infants than random control infants smoked during pregnancy (60% *v* 14%). These differences were significant in the multivariable analysis regardless of control group compared. The significance of covering the infant's head, postnatal exposure to tobacco smoke, dummy use, and sleeping in the side position has diminished.

Bias, confounding, and other reasons for caution

The noticeable reduction in SIDS rates and the small number of deaths from SIDS in this study and the difficulty in recruiting control families may limit the interpretation of the results. Our high risk control group was important as not only were their characteristics more similar to those of the families with SIDS infants but they also had more in common with deprived families.

We did not interview the control families at weekends, when alcohol and drug use may be more common. However, an analysis restricted to deaths and reference sleeps occurring only during the week suggests the combined effect of cosleeping and alcohol or drug use was still highly significant.

Generalisability to other populations

The south west region is representative of others in the United Kingdom—a predominantly white population and mix of rural and urban areas—and the increased deprivation among SIDS families reflects a similar trend in other populations. The prevalence of cosleeping differs between different cultures.

Study funding/potential competing interests

This study was funded by grants from the Foundation for the Study of Infant Deaths, Babes in Arms, and the Charitable Trust of University Hospitals Bristol. We have no competing interests.

MULTIVARIABLE LOGISTIC REGRESSION TO TEST FOR INTERACTION BETWEEN COSLEEPING AND ALCOHOL AND DRUG USE BY PARENTS OF STUDY INFANTS

	Model 1: SIDS infants (n=74) *v* random controls (n=86)*		Model 2: SIDS infants (n=74) *v* high risk controls (n=81)†	
Factors	Odds ratio (95% CI)	P value	Odds ratio (95% CI)	P value
Infant coslept in parental bed or on sofa for last sleep	5.41 (1.12 to 26.17)	0.04	5.23 (1.37 to 19.91)	0.02
Parental use of alcohol (>2 units) or drugs before last sleep	0.52 (0.10 to 2.72)	0.44	0.69 (0.16 to 3.00)	0.62
Interaction between cosleeping and alcohol or drug use	53.26 (4.07 to 696.96)	0.002	11.76 (1.40 to 99.83)	0.02

Both models adjusted for age, day or night sleep, maternal smoking during pregnancy, infant found sleeping prone, swaddling, gestational age, not sharing room, and fair or poor health in last 24 hours
*Also adjusted for maternal education
†Also adjusted for number of live births, young maternal age, poor socioeconomic status, and infant placed on pillow