The effects of postnatal health education for mothers on infant care and family planning practices in Nepal: a randomised controlled trial

Alison Bolam, Dharma S Manandhar, Purna Shrestha, Matthew Ellis, Anthony M de L Costello

Abstract

Objectives: To evaluate impact of postnatal health education for mothers on infant care and postnatal family planning practices in Nepal.
Design: Randomised controlled trial with community follow up at 3 and 6 months post partum by interview. Initial household survey of study areas to identify all pregnant women to facilitate follow up.
Setting: Main maternity hospital in Kathmandu, Nepal. Follow up in urban Kathmandu and a periurban area southwest of the city.
Subjects: 540 mothers randomly allocated to one of four groups: health education immediately after birth and three months later (group A), at birth only (group B), at three months only (group C), or none (group D).
Interventions: Structured baseline household questionnaire; 20 minute, one to one health education at birth and three months later.
Main outcome measures: Duration of exclusive breast feeding, appropriate immunisation of infant, knowledge of oral rehydration solution and need to continue breast feeding in diarrhoea, knowledge of infant signs suggesting pneumonia, uptake of postnatal family planning.
Results: Mothers in groups A and B (received health education at birth) were slightly more likely to use contraception at six months after birth compared with mothers in groups C and D (no health education at birth) (odds ratio 1.62, 95% confidence interval 1.06 to 2.5). There were no other significant differences between groups with regards to infant feeding, infant care, or immunisation.
Conclusions: Our findings suggest that the recommended practice of individual health education for postnatal mothers in poor communities has no impact on infant feeding, care, or immunisation, although uptake of family planning may be slightly enhanced.

Introduction

The rational approach to health promotion—that information given by health workers during clinic based or community based contacts will bring about a change in health behaviour—is still an integral part of primary healthcare strategies.[1 2] In practice, opportunities for one to one health education are given low priority by busy health workers. A survey of perinatal services across India reported that opportunities to give health education messages to mothers in the community were invariably missed.[3]

The effectiveness of health education has also been questioned. A recent review of over 500 articles about health education in developing countries found that only 11% described and evaluated actual attempts at health education. Of these, four described randomised studies and only three fulfilled the author's criteria for a rigorously designed evaluation.[4] In countries with few resources there is also a trade off between impact and sustainability. Interventions that are considered successful usually result from small scale, well resourced projects which cannot be reproduced on a large scale. One non-randomised evaluation of an initiative to encourage postnatal health education in a district hospital in Bihar, India, did show significant improvements in early breast feeding practices, although health education by the health workers was not maintained in the longer term.[5]

With increasing use of hospital maternity and immunisation services, especially in urban areas of the developing world,[6] perinatal contact with mothers represents an opportunity for health education about infant care and family planning. In developing countries 50-60% of infant deaths occur in the neonatal period,[7] and mortality from acute respiratory infections is highest in the first two months of life, when a mother's response to warning signs is crucial for survival. Failure to use postnatal contraception may also lead to an early repeat pregnancy, with attendant risks to maternal health.

In our study, the prospectively defined hypothesis was that one to one postnatal health education for mothers would positively affect their subsequent knowledge of and practices about infant care and family planning. Because an intervention should be feasible and sustainable on a large scale, education was restricted to a maximum of two contacts. Clinical objectives were to evaluate the impact of the intervention on uptake of immunisation, knowledge about and care of acute respiratory infections and diarrhoea in

Centre for International Child Health, Institute of Child Health, London WC1N 1EH
Alison Bolam, *research fellow*
Matthew Ellis, *research fellow*
Anthony M de L Costello, *reader in international child health*
Mother and Infant Research Activities (MIRA), Maternity Hospital, Kathmandu, c/o GPO Box 921, Kathmandu, Nepal
Dharma S Manandhar, *consultant paediatrician*
Purna Shrestha, *research assistant*
Correspondence to: Dr Costello a.costello@ich.ucl.ac.uk

BMJ 1998;316:805–11

infants, the duration of exclusive breast feeding, infant growth, and use of postnatal family planning services.

Subjects and methods

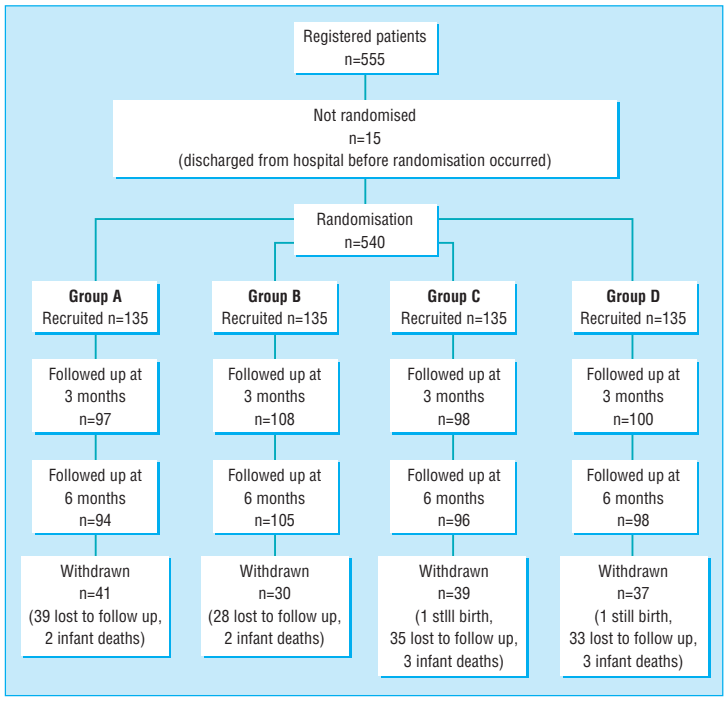
The study was conducted from November 1994 to May 1996. Oral consent from mothers for inclusion in the study was obtained before assignment. The study received ethical approval from the Nepal Health Research Council.

Setting

Nepal is one of the least developed countries in the world, with an infant mortality of 98/1000 live births, maternal mortality of 1500/100 000 live births, 26% adult literacy, and a prevalence of contraceptive use of 23%.[8] The estimated population of Kathmandu municipality is 500 000, with an annual urban growth rate of 7.4%.[9] Prasuti Griha is the main government funded maternity hospital in Kathmandu, with 250 beds, 15 000 deliveries annually, and outpatient services for the local urban and surrounding populations. As there are no formal addresses in Kathmandu, a house to house survey of two communities was conducted before the study. Kirtipur is a periurban area 5 km south west of the hospital that contains 3663 households with a total population of 21 368. It is a settled community of mainly wage labourers and farmers. Kalimati is an urban area of central Kathmandu situated 2 km from Prasuti Griha and containing 2467 households with a total population of 13 875. This is a mixed community of long term residents and recent migrants.

Eligibility

All pregnant women admitted to Prasuti Griha hospital for delivery residing in these two communities were eligible for entry to the trial. Two mothers entered into the trial whose deliveries resulted in a stillbirth were withdrawn from the trial and received neither the intervention nor follow up.



Flow chart showing progress of mothers through randomised trial

Protocol

The health education intervention was developed with hospital staff in collaboration with consultants experienced in health education and women's development. Three female health educators, two midwives, and one community health worker were trained to give the health education. All were fluent in the two local languages, Nepali and Newari, and conducted the education intervention in the appropriate language. The health education session lasted about 20 minutes and was designed to be interactive and supportive rather than prescriptive in style. It was tested beforehand with 20 mothers, and modifications were made in the light of this experience. The health educators were monitored weekly during the trial by two principal investigators to check the quality of the intervention with regards to the content and the style of delivery, especially the level of interaction, and constructive critical feedback was given.

The first education session was conducted in a quiet room before discharge from the hospital. On average, seven mothers were enrolled in the trial each week from 250-300 admissions to the postnatal wards, so the risks of contamination (mothers in different groups sharing information) were negligible given that mothers were seen individually for the education intervention. The second education session was conducted in the mothers' home three months after delivery (mean length of time 14.1 (SD 2.4) weeks).

Although the health education given at birth and three months covered broadly the same areas, more emphasis was placed on the importance of exclusive breast feeding in the first session and on the need for family planning in the second session. The topics covered were infant feeding, treatment of diarrhoea, recognition of and response to symptoms suggesting acute respiratory infection in young infants, the importance of immunisation, and the importance of contraception after the puerperium. For each topic, the mothers were initially asked questions (such as, "How are you planning to feed your baby?" and "How would you know if your baby had pneumonia?") and given time to respond to encourage interaction. A discussion would follow depending on the response. For each topic, the discussion led on to the health educator giving key messages (see box) illustrated with large pictures on a cloth flip chart developed by local artists from health materials supplied by Unicef.

At the end of each session the health educator repeated the key messages covered and asked the mother if she had any other questions.

Outcome measures

After birth, data were collected on the pregnancy; mode and outcome of delivery; and infant gestational age, birth weight, length, and head circumference. Infant weight was measured to the nearest 50 g with a Soenhle electric infant weighing scale. Infant length was measured to the nearest 0.5 cm with a Rollametre (Child Growth Foundation), and head circumference was measured to the nearest 0.1 cm with a tape meas-

Key messages given by health educators

Infant feeding—The advantages of breast feeding (it is clean, nutritious, prevents infection, helps family planning); the dangers of bottle feeding (risk of diarrhoea, reduction in mother's milk supply); how to increase the supply of breast milk (maternal diet, fluids, feeding soon after birth)
Diarrhoea—The dangers of diarrhoea; what to do if the infant develops diarrhoea (continue breast feeding, give oral rehydration solution, go to a health centre or doctor if diarrhoea persists)
Symptoms and response to acute respiratory infection—Visit a health worker if the infant develops cough, chest indrawing, fast breathing, or poor feeding
Immunisation—The importance of full immunisation; where to go for the first or subsequent injections
Family planning—The importance of restarting contraception no later than 8 weeks after birth; the location of the nearest family planning clinic; the choice of methods; the availability of sterilisation services at the hospital

ure. Gestational age was assessed by the Parkin method.[10]

Women were followed up at 3 and 6 months post partum in their homes, when data relating to our outcome measures were collected. Primary outcomes were the duration of exclusive breast feeding; mothers' knowledge of important signs of pneumonia and appropriate management of diarrhoea (mothers were asked: "How do you know if your baby with cough has pneumonia?" and, "If your baby has diarrhoea how must you care for him?"); uptake of immunisation; and use of postnatal family planning services. A secondary outcome was infant nutritional status.

Sample size

To assess the baseline situation we reviewed recent national survey data and conducted a pilot survey at the hospital outpatient clinic of 200 postnatal mothers—100 at birth and 100 at up to 4 months post partum.

Duration of exclusive breast feeding—The pilot survey showed 34% mothers were exclusively breast feeding at 4 months post partum. These were well motivated women attending the hospital postnatally for infant immunisation. Another study of breast feeding practices in Nepal showed that only 20% of infants in Kathmandu were exclusively breast fed by 4 months of age.[11] We hypothesised that 25% of mothers with no educational intervention would exclusively breast feed at 4 months, with an improvement to 40% in the intervention group: to detect this difference with 95% confidence limits and a power of 80%, we needed to enter 165 mothers into each group.

Infant nutritional status—If education helps to prolong the duration of exclusive breast feeding, nutritional outcome might be improved. Assuming infants of mothers receiving the intervention grew on average along the 50th centile for British infants, a difference of 300 g in weight at 6 months (3.8%) between the group receiving no education with the two groups receiving health education at birth would be detected with 95% confidence limits and a power of 80% with sample sizes of 131 and 262.

Mothers' knowledge of managing infant diarrhoea and acute respiratory infection—A Nepal national survey in 1994 found that only 37% of urban children with diarrhoea received oral rehydration solution.[12] In Pakistan few mothers spontaneously mentioned rapid breathing as a sign of pneumonia,[13] and in the Philippines only 22% of cases of severe acute respiratory infection were recognised as severe by the mothers.[14] We hypothesised that 40% of control mothers and 60% of intervention mothers would correctly describe the signs of pneumonia and how to manage diarrhoea, requiring 107 mothers in each group (95% confidence interval, power 80%).

Immunisation uptake—In 1991 in Nepal 74% of children aged 12 months had received three doses of diphtheria and pertussis vaccine and oral polio vaccine, and 81% were vaccinated against tuberculosis.[15] We hypothesised that 40% of control and 60% of intervention infants would be fully immunised by 6 months of age, requiring 107 mothers in each group (95% confidence interval, power 80%).

Family planning—In our pilot survey 20% of postnatal mothers were using a method of contraception at 4 months. Use of contraceptives by currently married women nationally was estimated at 14%,[15] but this can be assumed to be higher in urban areas. We hypothesised a 20% uptake of contraception in the

Table 1 Baseline details of 540 mothers and infants recruited for study. Values are numbers (percentages) unless stated otherwise

Variable	Intervention group*				All subjects (n=540)
	Group A (n=135)	Group B (n=135)	Group C (n=135)	Group D (n=135)	
Mothers					
Mean (SD) age (years) (n=538)	23.4 (4.0)	23.2 (4.3)	23.1 (4.2)	24.0 (3.4)	23.4 (4.0)
Parity (n=510):					
Primigravida	61 (50)	65 (50)	77 (60)	68 (53)	271 (53)
Multigravida	62 (50)	65 (50)	52 (40)	60 (47)	239 (47)
Area of living:					
Periurban	51 (37)	54 (40)	45 (33)	50 (38)	200 (37)
Urban	84 (63)	81 (60)	90 (67)	85 (62)	340 (63)
Woman's occupation (n=508):					
Housewife	102 (84)	106 (82)	111 (86)	102 (80)	421 (83)
Other	20 (16)	24 (18)	18 (14)	25 (20)	87 (17)
Woman's educational status (n=509):					
Illiterate	39 (32)	26 (20)	37 (29)	38 (30)	140 (28)
Literate up to primary level only	26 (21)	38 (30)	31 (15)	27 (21)	122 (24)
Secondary level and above	58 (48)	66 (50)	60 (47)	63 (49)	247 (49)
Husband's monthly income (n=500):					
Low (<$20)	6 (5)	10 (8)	6 (5)	3 (2)	25 (5)
Medium ($20-100)	102 (84)	96 (74)	99 (80)	105 (84)	402 (81)
High (>$100)	14 (12)	23 (18)	19 (15)	17 (14)	73 (15)
Delivery mode:					
Normal	122 (90)	118 (87)	111 (82)	105 (78)	456 (84)
Assisted vaginal or caesarean section	13 (10)	17 (13)	24 (18)	30 (22)	85 (16)
Infants					
Sex:					
Male	75 (55)	60 (45)	63 (47)	67 (49)	265 (49)
Female	60 (45)	75 (55)	72 (53)	68 (51)	275 (51)
Mean (SD) weight (kg) (n=533)	2.75 (0.4)	2.75 (0.4)	2.78 (0.5)	2.67 (0.5)	2.74 (0.5)
Mean (SD) length (cm) (n=528)	48.5 (2.7)	48.6 (2.8)	48.7 (3.3)	48.3 (3.4)	48.5 (3.1)
Mean (SD) head circumference (cm) (n=531)	33.6 (1.3)	33.8 (2.3)	33.8 (1.9)	33.9 (2.9)	33.8 (2.2)
Mean (SD) gestation (weeks) (n=514)	39.3 (2.2)	39.4 (1.7)	39.4 (1.8)	39.1 (2.4)	39.3 (2.0)

*Group A=health education given immediately after birth and 3 months later; group B=education given at birth only; group C=education given at 3 months only; group D=no education given.

Table 2 Outcomes recorded at 3 months post partum for 403 mothers and infants. Values are numbers (percentages) unless stated otherwise

Variable	Intervention group*: Groups A and B	Intervention group*: Groups C and D	Odds ratio (95% CI)	P value
Infant feeding practice:				
Exclusive breast feeding	120 (59)	117 (59)	1.00 (0.67 to 1.49)	1.00
Other	84 (41)	82 (41)		
Mean (SD) weight (kg)	5.7 (0.85)	5.6 (0.9)	Mean difference 0.100 (–0.072 to 0.271)	0.25
Mean (SD) length (cm)	58.4 (4.6)	58.5 (6.2)	Mean difference –0.041 (–1.11 to 1.03)	0.94
Mean (SD) head circumference (cm)	40.9 (2.7)	40.7 (4.1)	Mean difference 0.217 (–0.469 to 0.902)	0.53
Immunisation:				
Appropriate†	179 (87)	169 (85)	1.18 (0.67 to 2.08)	0.66
Other	26 (13)	29 (15)		
Knowledge of signs of pneumonia:				
Indrawing:				
Yes	54 (26)	44 (22)	1.25 (0.79 to 1.97)	0.35
No	151 (74)	154 (78)		
Tachypnoea:				
Yes	107 (52)	84 (42)	1.48 (1.00 to 2.19)	0.06
No	98 (48)	114 (58)		
In case of diarrhoea knows to:				
Continue breast feeding:				
Yes	94 (46)	94 (47)	0.95 (0.64 to 1.40)	0.84
No	111 (54)	105 (53)		
Give oral rehydration solution:				
Yes	177 (86)	165 (83)	1.26 (0.74 to 2.17)	0.41
No	28 (14)	33 (17)		
Any contraceptive used for family planning:				
No	163 (80)	171 (86)	1.49 (0.87 to 2.53)	0.14
Yes	40 (20)	28 (14)		

*Groups A and B=health education given at birth; groups C and D=no health education at birth.
†BCG plus at least two doses of diphtheria and pertussis vaccine and oral polio vaccine.

control group by 6 months post partum and an improvement to 33% in the intervention group, requiring 195 mothers in each group (95% confidence interval, power 80%).

Using these figures, we enrolled 540 subjects in order to compare four subgroups: mothers receiving health education immediately after birth and at 3 months post partum (group A), health education at birth only (group B), health education at three months only (group C), or no health education at all (group D). For outcomes at three months, we combined groups A and B as the intervention group and C and D as the control group. For the outcomes at 6 months, the groups were compared individually.

Randomisation and blinding

The unit of randomisation was the individual mother. Restricted randomisation was used in blocks of 20, each block consisting of a random ordering of the numbers 0-19. Numbers 0-4, 5-9, 10-14, and 15-19 were assigned to groups A to D respectively. The details of allocation to groups for consecutively recruited mothers were in sealed envelopes. Timing of assignment was when a mother was identified by the research team either in labour or shortly after delivery. A member of the research team checked the hospital admission register at least twice each day between 7 am and 8 pm. The generator of the assignment was not involved in the execution of the allocation. There were no prospectively defined rules for stopping the trial.

Clearly, the mothers recruited and the health educators were not blind to the assignment of mothers to different groups. The outcome assessors were always blind to the assignment at both the 3 and 6 month follow up visits. Staff who were involved in data collection at the 3 month follow up were not involved in data collection at 6 months. The data analysts were not blind to the coding of the groups.

Statistical analysis

To estimate the effect of the trial intervention between the groups we measured the mean differences, 95% confidence intervals, and P values for continuous data, and the odds ratios, 95% confidence intervals, and P values for categorical data. We used the Mantel-Haenszel test to check for heterogeneity of categorical data, giving χ^2 and P values, and analysis of variance for continuous data, giving F values and P values. We analysed data on an intention to treat basis in which we compared intervention and control groups irrespective of the quality of the education intervention. For statistical analysis, we used computer software Statview version 4.0 and Stata version 5.0.

Results

Subjects

The figure shows the details of participant flow and follow up. We recruited 540 mothers, 135 to each of the four groups, and followed up 403 (75%) to 3 months post partum and 393 (73%) to 6 months. The main reason for loss to follow up was the mother moving back to her parental home as part of cultural tradition. Table 1 shows the baseline characteristics of the mothers and infants.

Mortality—There were no maternal deaths, two stillbirths, and 10 infant deaths. Mothers whose infants were stillborn were withdrawn from the study. All the infant deaths occurred in the neonatal period: two occurred in group A, two in group B, three in group C, and three in group D. Seven of these infants were born prematurely and had a birth weight less than 2.5 kg, two had severe congenital abnormalities, and one died from acute respiratory infection at home at 4 weeks of age.

Outcome at 3 months

Table 2 shows the outcomes at 3 months post partum. We compared mothers in groups A and B, who received health education at birth, with those in groups C and D, who received none. Mothers in groups A and B were slightly more likely to report tachypnoea as a sign of acute respiratory infection, but this did not quite reach statistical significance (odds ratio 1.48, 95% confidence interval 1.00 to 2.19, P = 0.06). Also, 20% of mothers in groups A and B were using contraception compared with only 14% of those in groups C and D, but this difference was not significant. There were no differences for the other outcomes.

Immunisation coverage was higher than we had hypothesised for both groups (85% in groups C and D, 87% in groups A and B): our sample size would have detected an increase to 93% coverage in groups A and B at 5% significance (one sided test) and 78% power .

Table 3 Outcomes recorded at 6 months post partum for 393 mothers and infants. Values are numbers (percentages) unless stated otherwise

Variable	Intervention group*				Groups A and B *v* C and D		Groups A and C *v* B and D		Outcome by health education at birth stratified by health education at 3 months
	Group A	Group B	Group C	Group D	Odds ratio (95% CI)	P value	Odds ratio (95% CI)	P value	
Duration of exclusive breast feeding (n=390):									
⩾5 months	31 (33)	25 (24)	27 (29)	27 (28)	1.01 (0.65 to 1.56)	1.00	1.29 (0.83 to 2.0)	0.31	Mantel-Haenszel test: χ^2=0.72, P=0.395
<5 months	63 (67)	79 (76)	67 (71)	71 (72)					
Mean (SD) weight (kg)	7.2 (0.9)	7.3 (1.0)	7.2 (0.93)	7.2 (1.1)	Mean difference 0.028 (–0.169 to 0.225)	0.78	Mean difference –0.06 (–0.255 to 0.139)	0.56	Analysis of variance: Education at 3 months *F*=0.085, P=0.77 No education *F*=0.01, P=0.92
Mean (SD) length (cm)	62.4 (7.4)	63.2 (4.0)	62.3 (7.0)	62.9 (4.5)	Mean difference 0.223 (–0.95 to 1.39)	0.71	Mean difference –0.7 (–1.87 to 0.47)	0.24	Analysis of variance: Education at 3 months *F*=0.006, P=0.94 No education *F*=0.28, P=0.60
Mean (SD) head circumference (cm)	42.5 (1.5)	42.4 (2.0)	42.5 (1.4)	42.6 (2.1)	Mean difference –0.06 (–0.41 to 0.29)	0.74	Mean difference –0.32 (–0.39 to 0.32)	0.86	Analysis of variance: Education at 3 months *F*=0.008, P=0.93 No education *F*=0.23, P=0.63
Immunisation:									
Appropriate†	90 (95)	100 (96)	90 (93)	91 (94)	1.52 (0.65 to 3.55)	0.39	0.79 (0.34 to 1.82)	0.66	Mantel-Haenszel test: χ^2=0.033, P=0.855
Other	5 (5)	4 (4)	7 (7)	6 (6)					
Knowledge of signs of pneumonia:									
Indrawing:									
Yes	27 (28)	26 (25)	26 (27)	20 (21)	1.17 (0.75 to 1.84)	0.56	1.28 (0.82 to 2.02)	0.30	Mantel-Haenszel test: χ^2=0.13, P=0.72
No	68 (72)	78 (75)	71 (73)	77 (79)					
Tachypnoea:									
Yes	51 (54)	56 (54)	56 (58)	44 (45)	1.09 (0.74 to 1.62)	0.69	1.27 (0.86 to 1.89)	0.27	Mantel-Haenszel test: χ^2=1.54, P=0.21
No	44 (46)	48 (46)	41 (42)	53 (55)					
In case of diarrhoea knows to:									
Continue breast feeding:									
Yes	50 (53)	51 (49)	49 (51)	48 (49)	1.03 (0.69 to 1.53)	0.92	1.10 (0.74 to 1.63)	0.69	Mantel-Haenszel test: χ^2=0.064, P=0.799
No	45 (47)	53 (51)	48 (49)	49 (51)					
Give oral rehydration solution:									
Yes	89 (94)	96 (92)	94 (97)	90 (93)	0.72 (0.32 to 1.63)	0.53	1.64 (0.71 to 3.76)	0.3	Mantel-Haenszel test: χ^2=0.568, P=0.45
No	6 (6)	8 (8)	3 (3)	7 (7)					
Any contraceptive used for family planning:									
No	62 (65)	64 (62)	72 (74)	71 (73)	1.62 (1.06 to 2.50)	0.03	0.86 (0.58 to 1.35)	0.59	Mantel-Haenszel test: χ^2=0.06, P=0.81
Yes	33 (35)	40 (38)	25 (26)	26 (27)					

*Group A=health education given immediately after birth and at 3 months later; group B=education given at birth only; group C=education given at 3 months only; group D=no education given.
†BCG plus three doses of diphtheria and pertussis vaccine and oral polio vaccine.

Outcome at 6 months

Table 3 shows the outcomes at 6 months post partum. We made two broad comparisons: groups A and B (health education at birth) compared with groups C and D (no health education at birth), and groups A and C (health education at 3 months) compared with groups B and D (no health education at 3 months). The only significant difference we observed for all outcomes was an increase in uptake of family planning at 6 months in groups A and B (odds ratio 1.62, 95% confidence interval 1.06 to 2.5). To test for interactions, we compared outcomes by health education at birth stratified by whether health education was given at 3 months post partum using tests for heterogeneity: we found no significant interactions.

Poststudy calculations of the power of our study to detect a significant, one sided difference in exclusive breast feeding between groups (based on our hypothesis of 25% in mothers given no health education and 40% in those given education) were 67% (comparing group A with group D) and 84% (comparing groups A, B, and C with group D).

Discussion

This trial in Nepal has shown that a health education intervention (one to one counselling of mothers by health educators) given on two occasions, immediately after delivery and 3 months later, had no significant impact on the mothers' knowledge and practices of child care or infant health outcomes, but there was a slight improvement in uptake of family planning at 6 months after birth. Given the higher than expected level of immunisation in all groups, we cannot rule out the possibility that health education may have had an impact in situations where coverage is lower.

Trial design

Our study included only women who chose institutional delivery. Whether mothers who gave birth at home would benefit from health education more than those who gave birth at hospital is questionable, but it is difficult to target mothers delivering at home and to conduct a trial of intervention in the home.

The overall lack of impact on practices in infant care might also be explained by the length and frequency of the intervention. Our study deliberately

involved a maximum of only two contacts with each mother in an attempt to evaluate a less intensive, more sustainable intervention. There is some evidence that health education at an individual level has an impact if messages are repeated frequently to patients,[4] but multiple contact with patients in the community is difficult to sustain in a resource poor country such as Nepal. Pilot studies that report success are usually from well funded, small scale, non-government projects.[16]

A combination of antenatal and perinatal contacts might be more successful. In this trial 88% of women had attended one or more antenatal clinic appointments, at which only 3% had received any health education. Follow up rates for the trial were less than ideal (75% at 3 months and 73% at 6 months) but reasonable for a trial conducted in difficult field conditions, where mothers often return to their parental home postnatally.

Evaluation of health education interventions

Recommendations for the design of health education interventions and the importance of including evaluation in health education programmes have been widely reported.[4 17–20] For example, the American Public Health Association stated that "from the outset, a health promotion program should be organised, planned and implemented in such a way that its operation and effects can be evaluated."[17] In practice, however, evaluation is rare. In a review of health education in developing countries spanning 10 years, only 11% of published articles described and evaluated the health education programme.[4] Most of the evaluations were methodologically unsound so firm conclusions could not be drawn about the overall efficacy of health education. Randomised controlled studies, the ideal design, have rarely been reported from developing countries: only four of the studies reviewed by Loevinsohn used a randomised controlled design, and one of these failed to meet other of his criteria for an adequate study design.[4]

Health education by health workers is still seen as an important part of primary health care despite this lack of evidence of efficacy. Training of health and field workers to convey messages, and the development of health education materials, consumes a substantial proportion of health budgets in resource poor countries. Our negative findings suggest that much of this investment may be ineffective. Social cognitive theory, by contrast, suggests that experience from interactions within family, peer groups, or communities, rather than information per se, is the key to successful health promotion.[21 22]

It might be argued that a postnatal health intervention would be more effective if it focuses on only one outcome. The small but significant increase in contraceptive use at 6 months post partum by the mothers receiving health education immediately after birth might have been even greater if this was the only subject discussed. This requires further evaluation because postnatal family planning and birth spacing have health benefits for both mothers and infants. It might also be argued that mothers in Nepal do not perceive many health workers as purveyors of credible knowledge about motherhood. In our study we deliberately selected health educators who were able to gain the respect of mothers through their experience as midwives or community health workers, but who were also able to put mothers at ease during the education session.

Key messages

- Health education is widely promoted in primary care, but there have been few rigorous evaluations of its impact, especially in developing countries
- A randomised controlled trial of postnatal individual health education for mothers given by trained female health workers showed no significant impact on maternal knowledge and practices of child care or on infant health outcomes, but there was a small improvement in uptake of family planning at six months after birth
- The efficacy of health education interventions that rely solely on giving people information to bring about a change in health behaviour is unproved; interventions should be evaluated before being implemented on a large scale
- Alternative strategies for health promotion in developing countries such as interactions within families, peer groups, or communities may be more effective but are costly and difficult to implement on a large scale

Conclusions

Our results indicate the need for further, well designed evaluations of health education interventions that are randomised and controlled, provide a clear definition of aims, and present pre-intervention and post-intervention data for carefully defined outcome measures. Future evaluations of education interventions also need to explore, through qualitative research, the understanding of the recipients and their reaction to the messages. It might be that behaviour can be changed in response to simple messages repeated frequently in many forums, but in developing countries there will a trade off between efficacy and cost: repeated home visits by friendly health workers may not be feasible on a large scale. It might also be the case that the desired changes in behaviour are not realistic for the individual or community because of economic, social, and cultural barriers. Interventions aimed at women must take into account their heavy workload in the home and field and their degree of influence within the household on decisions about child care, family planning, and health seeking behaviour.

We thank Dr Kasturi Malla for helpful discussions and the staff of Prasuti Griha and MIRA, Nepal, for their support. We are grateful to Diana Gibb, Ruth Gilbert, and Sally Kerry for advice on trial design and statistical analysis.

Contributors: AB coordinated the formulation of the trial design and protocol, supervised the trial implementation, and participated in data collection and analysis and writing of the paper. DSM helped to formulate the initial study hypothesis, trial design, and protocol; assisted with study implementation; and reviewed drafts of the paper. PS contributed to the study design, participated in data collection, and reviewed drafts of the paper. ME contributed to formulation of the trial design and protocol, and participated in data analysis and writing of the paper. AMdeLC conceived of the study hypothesis, developed the trial design and protocol, and contributed to data analysis

and writing of the paper. AB and AMdeLC are guarantors for the paper.

Funding: Research grant from Britain's Department for International Development.

Conflict of interest: None.

1 World Health Organisation. *The Alma Ata declaration*. Geneva: WHO, 1978.
2 World Bank. *Investing in health. The world development report*. Oxford: Oxford University Press, 1993.
3 Bhargava SK, Singh KK, Saxena BN, eds. *A national collaborative study of identification of high risk families, mothers and outcome of their offsprings with particular reference to the problem of maternal nutrition, low birth weight, perinatal and infant morbidity and mortality in rural and urban slum communities*. New Delhi: Indian Council of Medical Research, 1990.
4 Loevinsohn BP. Health education in developing countries: a methodological review of published articles. *Int J Epidemiol* 1990;19: 788-94.
5 Prasad B, Costello AM de L. Impact and sustainability of a "baby friendly" health education intervention at a district hospital in Bihar, India. *BMJ* 1995;310:621-3.
6 Bolam A, Manandhar DS, Shrestha P, Manandhar B, Ellis M, Costello AM. Maternity care utilisation in the Kathmandu Valley: a community based study. *J Nepal Med Assoc* 1997;35:122-9.
7 Ashworth A, Waterlow JC. Infant mortality in developing countries. *Arch Dis Child* 1982;57:882-4.
8 *Human Development Report. Published for the United Nations Development Programme*. Oxford: Oxford University Press, 1996.
9 World Bank. *World development report 1990*. Washington DC: World Bank, 1990.
10 Parkin JM, Hey EN, Clowes JS. Rapid assessment of gestational age in the newborn infant. *Arch Dis Child* 1976;51:259.
11 Neupane S. *A study on breast feeding status in rural and urban areas of central development region, Nepal*. Kathmandu: Unicef, 1992.
12 *Nepal multiple indicator surveillance: preliminary report*. Kathmandu: National Planning Commission, His Majesty's Government of Nepal, 1995.
13 Malik Kundi MZ, Anjum M, Mull DS, Mull JD. Maternal perceptions of pneumonia and pneumonia signs in Pakistani children. *Soc Sci Med* 1993;37:649-60.
14 Tupasi TE, Miguel CA, Tallo VL, Bagasao TMP, Natividad JN, Valencia LB, et al. Child care practices of mothers: implications for intervention in acute respiratory infections. *Ann Tropical Paediatrics* 1989;9:82-8.
15 United Nations Children's Fund. *The state of the world's children 1993*. Oxford: Oxford University Press, 1993.
16 Kafatos A, Tsitoura S, Pantelakis S, Doxiadis S. Maternal and infant health education in a rural Greek community. *Hygie* 1991;10:32-7.
17 American Public Health Association Work Group on Health Promotion/ Disease Prevention. Criteria for the development of health promotion and education programs. *Am J Public Health* 1987;77:89-92.
18 Hubley JH. Barriers to health education in developing countries. *Health Educ Res* 1986;1:233-45 .
19 Oakley A, Fullerton D, Holland J, Arnold S, France-Dawson M, Kelley P, et al. Sexual health education interventions for young people: a methodological review *BMJ* 1995;310:158-62.
20 Gyorkos TW, Tannenbaum TN, Abrahamowicz M, Oxman AD, Scott EA, Millson ME, et al. An approach to the development of practice guidelines for community health interventions. *Can J Public Health* 1994;85:S8-13.
21 Maibach E, Murphy DA. Self-efficacy in health promotion research and practice: conceptualisation and measurement. *Health Educ Res* 1995;10:37-50.
22 Bandura A. Self-efficacy: toward a unifying theory of behavioural change. *Psychol Rev* 1977;84:191-215.

(Accepted 6 February 1998)

Meta-analysis of short term low dose prednisolone versus placebo and non-steroidal anti-inflammatory drugs in rheumatoid arthritis

Peter C Gøtzsche, Helle Krogh Johansen

Abstract

Objective: To determine whether short term, oral low dose prednisolone (≤15 mg daily) is superior to placebo and non-steroidal anti-inflammatory drugs in patients with rheumatoid arthritis.

Design: Meta-analysis of randomised trials of oral corticosteroids compared with placebo or a non-steroidal anti-inflammatory drug.

Setting: Trials conducted anywhere in the world.

Subjects: Patients with rheumatoid arthritis.

Main outcome measures: Joint tenderness, pain, and grip strength. Outcomes measured on different scales were combined by using the standardised effect size (difference in effect divided by SD of the measurements).

Results: Ten studies were included in the meta-analysis. Prednisolone had a marked effect over placebo on joint tenderness (standardised effect size 1.31; 95% confidence interval 0.78 to 1.83), pain (1.75; 0.87 to 2.64), and grip strength (0.41; 0.13 to 0.69). Measured in the original units the differences were 12 (6 to 18) tender joints and 22 mm Hg (5 mm Hg to 40 mm Hg) for grip strength. Prednisolone also had a greater effect than non-steroidal anti-inflammatory drugs on joint tenderness (0.63; 0.11 to 1.16) and pain (1.25; 0.26 to 2.24), whereas the difference in grip strength was not significant (0.31; – 0.02 to 0.64). Measured in the original units the differences were 9 (5 to 12) tender joints and 12 mm Hg (– 6 mm Hg to 31 mm Hg). The risk of adverse effects during moderate and long term use seemed acceptable.

Conclusion: Prednisolone in low doses (≤15 mg daily) may be used intermittently in patients with rheumatoid arthritis, particularly if the disease cannot be controlled by other means.

Introduction

Corticosteroids were first shown to be effective in patients with rheumatoid arthritis in 1949 in an uncontrolled study.[1] In 1959, a two year randomised trial showed that an initial dose of prednisolone 20 mg daily was significantly superior to aspirin 6 g daily.[2] Important adverse effects were also noted, however, and the authors concluded that the highest acceptable dose for long term treatment was probably in the region of 10 mg daily.

Corticosteroids have received renewed interest in recent years because of their possible beneficial effect on radiological progression.[3] Tendencies towards such an effect were noted both in the early trials and in a recent report.[4]

These findings are interesting, but oral corticosteroids are still being used mainly for their symptomatic effect—for example, for acute exacerbations of rheumatoid arthritis and as "bridge therapy" before slow acting drugs have taken effect.[5] The effect of low doses

Editorial by Dennison and Cooper

Nordic Cochrane Centre, Rigshospitalet, Department 7112, Tagensvej 18 B, DK-2200 Copenhagen N, Denmark

Peter C Gøtzsche, *director*

Helle Krogh Johansen, *senior researcher*

Correspondence to: Dr Gøtzsche p.c.gotzsche@cochrane.dk

BMJ 1998;316:811–8

has been variable, however, and was questioned as late as 1995 when the most recent trial of low dose steroids was published.[6] We therefore performed a systematic review of randomised trials that compared corticosteroids, given at a dose equivalent to no more than 15 mg prednisolone daily, with placebo or with non-steroidal anti-inflammatory drugs. Our review is limited to the short term effect—that is, as recorded within the first weeks of treatment. In an analysis of the adverse effects of steroids, however, we also included long term trials and matched cohort studies.

Methods

All randomised studies that compared an oral corticosteroid with placebo or a non-steroidal anti-inflammatory drug in patients with rheumatoid arthritis were eligible if they reported clinical outcomes within 1 month after the start of treatment. When there were data from several visits, the data that came closest to 1 week of treatment were used for the analyses. We excluded studies with high dose steroids (exceeding an equivalent of 15 mg prednisolone daily); studies of combination treatments—for instance, of a steroid and a non-steroidal anti-inflammatory drug; and studies that used quasi-randomisation methods, such as allocation by date of admission or by toss of a coin (no such studies were actually found). The outcome variables were joint tenderness (usually Ritchie's joint index), pain, and grip strength.

Medline was searched from 1966 onwards and most recently updated in September 1997. We used the *Explode* option (which searches for a broad term plus related narrower items) for "glucocorticoids" *or* "glucocorticoids, -synthetic" (for all subheadings) combined with *Explode* "arthritis-rheumatoid" (for all subheadings) and with "placebos" *or* "comparative study" in MeSH. The reference lists were scanned for additional trials, and an archive in possession of one of the authors was searched. As most of the retrieved trials were very old and the steroid drugs were non-proprietary ones authors and companies were not asked about possible unpublished studies. We did not handsearch journals for relevant trials as this work is already being organised by the Cochrane Collaboration for all medical journals, including specialist rheumatological journals. The results of these handsearches are made available in the Cochrane Controlled Trials Register in *The Cochrane Library*,[7] which we searched with prednisolone and prednisone as text words combined with rheumatoid.

Decisions on which trials to include were taken independently by two observers based only on the methods sections of the trials; disagreements were resolved by discussion. Details on the nature and dose of treatments, number of randomised patients, the randomisation and blinding procedures, and exclusions after randomisation were noted. When an outcome was measured on the same scale in all trials we calculated the weighted mean difference as the summary estimate for the effect. As the outcomes were often measured on different scales, however, even when they referred to the same quality—for example, tender joints—we also calculated standardised effect measures.[8] With this method the difference in effect between two treatments is divided by the standard deviation of the measurements. By that transformation the effect measures become dimensionless, and outcomes from trials which have used different scales may therefore often be combined. As an example, the tender joint count may be recorded either as the number of tender joints or as Ritchie's index, in which each joint is scored on a scale from 0 to 3 for pain on firm palpation and the scores added. Often the two types of counts will give similar values, but if the patients have very severe disease Ritchie's index may be higher. The standard deviation will then also be higher, however, and by dividing the counts with their standard deviations (for example, of the baseline measurements) the effect sizes will be of the same magnitude.

The random effects model[9] was used if $P < 0.10$ for the test of heterogeneity; otherwise a fixed effects analysis was performed. As data from crossover trials were reported in only summary form, as if they had been generated from a group comparative trial, we analysed them accordingly. We therefore assumed that no important carryover effects had occurred.

Table 1 General characteristics of studies included in meta-analysis of low dose prednisolone in treatment of rheumatoid arthritis

Study	Design	Study drugs		Length of treatment (days)
		Prednisolone	Control	
Berry 1974[33]	Crossover	15 mg	Placebo	7
Boardman 1967[34]*	Crossover	7.5 mg	Placebo	7
Böhm 1967[35 36]	Crossover	2.5 mg	Placebo	8
Dick 1970[37]	Crossover	10 mg	Placebo; ibuprofen 1200 mg; aspirin 4 g†	7
Gestel 1995[6 32]	Parallel	10 mg	Placebo	7‡
Jasani 1968[38]	Crossover	15 mg	Placebo; ibuprofen 750 mg; aspirin 5 g†	7
Lee 1973[39]	Crossover	15 mg	Placebo; aspirin 5 g	7
Lee 1973[40 41]	Parallel	15 mg	Placebo; aspirin 3.9 g	14
Lee 1974[42]	Crossover	10 mg	Placebo; sodium salicylate 4 g	7
Stenberg 1992[43]	Crossover	3 mg	Placebo	5§

*We included two patients in analysis (excluded by authors because of too little difference in joint size) by assuming that difference in grip strength was 0.
†Average of ibuprofen and aspirin used in analysis.
‡One week data provided by authors.
§Each flare treated for 5 days; three randomised patients who were excluded because of poor response to prednisolone in introductory test period included in analysis by assuming that difference between prednisolone and placebo was 0.

Results

Twenty eight randomised trials were initially identified, several of which had been published more than once. Eighteen trials were excluded for various reasons.[2 4 5 10–31] Nine trials did not fulfil the inclusion criteria for the meta-analysis: five had studied combinations of drugs[10 17–19 27 31]; two used too high a dose[2 20–23]; in one, 4 mg methylprednisolone was given to all the patients in the placebo group[28]; and one concerned patients with juvenile rheumatoid arthritis (this trial found prednisolone to be significantly better than placebo).[24]

The other nine excluded studies were potentially eligible for the meta-analysis. However, one was a five way crossover trial with a grossly unbalanced design—for instance, placebo was given to 9, 13, 3, 6, and 6 patients during weeks 1, 2, 3, 4, and 5, respectively.[12] Because of regression towards the mean we found it inappropriate to include this trial. Another trial was also unbalanced as the steroid group was kept mobile

Study	Experimental treatment: No of subjects	Experimental treatment: Mean (SD)	Control treatment: No of subjects	Control treatment: Mean (SD)	Weight (%)	Standardised mean difference (95% CI)
Joint tenderness (Ritchie's index)						
Berry 1974[33]	12	13.0 (11.0)	12	23.7 (11.0)	13.4	-0.939 (-1.790 to -0.088)
Dick 1970[37]	24	17.6 (8.0	24	40.7 (13.0)	15.0	-2.105 (-2.822 to -1.389)
Gestel 1995[6 32]	20	10.8 (4.7)	20	16.3 (7.7)	15.7	-0.845 (-1.495 to -0.195)
Jasani 1968[38]	9	16.2 (8.7)	9	38.1 (12.8)	10.3	-1.906 (-3.068 to -0.744)
Lee 1973[39]	21	30.5 (16.5)	21	41.4 (19.8)	16.1	-0.587 (-1.206 to 0.032)
Lee 1974[42]	18	14.6 (12.4)	18	26.4 (15.1)	15.3	-0.835 (-1.520 to -0.151)
Stenberg 1992[43]	21	6.3 (1.7)	21	11.1 (2.5)	14.2	-2.203 (-2.985 to -1.421)
Total	125		125		100.0	-1.305 (-1.828 to -0.782)
χ^2 = 20.21, df = 6, Z = 4.89						
Pain (ranking scale, visual analogue scale, or composite)						
Böhm 1967[35 36]	20	2.15 (0.99)	20	2.60 (0.94)	17.6	-0.457 (-1.086 to 0.172)
Dick 1970[37]	24	0.46 (0.59)	24	2.83 (0.29)	14.7	-4.661 (-5.787 to -3.536)
Gestel 1995[6 32]	20	35.6 (16.2)	20	58.3 (21.2)	17.4	-1.179 (-1.856 to -0.502)
Jasani 1968[38]	9	5.7 (5.7)	9	25.1 (14.6)	14.8	-1.667 (-2.778 to -0.557)
Lee 1973[40 41]	45	2.56 (0.83)	41	3.47 (0.83)	18.4	-1.087 (-1.541 to -0.632)
Stenberg 1992[43]	21	23.5 (5.9)	21	39.7 (9.90)	17.0	-1.950 (-2.697 to -1.204)
Total	139		135		100.0	-1.752 (-2.638 to -0.865)
χ^2 = 45.32, df = 5, Z = 3.87						
Grip strength (mm Hg)						
Boardman 1967[34]	13	372.0 (85.0)	13	299.0 (85.0)	11.7	0.832 (0.025 to 1.639)
Dick 1970[37]	24	213.0 (136.0)	24	149.0 (115.0)	23.0	0.500 (-0.075 to 1.075)
Gestel 1995[6 32]	20	191.0 (112.0)	20	160.0 (160.0)	19.7	0.220 (-0.402 to 0.842)
Jasani 1968[38]	8	356.0 (151.0)	8	267.0 (125.0)	7.5	0.607 (-0.402 to 1.617)
Lee 1973[39]	21	109.0 (47.0)	21	97.0 (47.0)	20.6	0.251 (-0.357 to 0.858)
Lee 1974[42]	18	73.1 (43.5)	18	59.2 (39.1)	17.6	0.329 (-0.330 to 0.987)
Total	104		104		100.0	0.410 (0.134 to 0.686)
χ^2 = 1.97, df = 5, Z = 2.91						

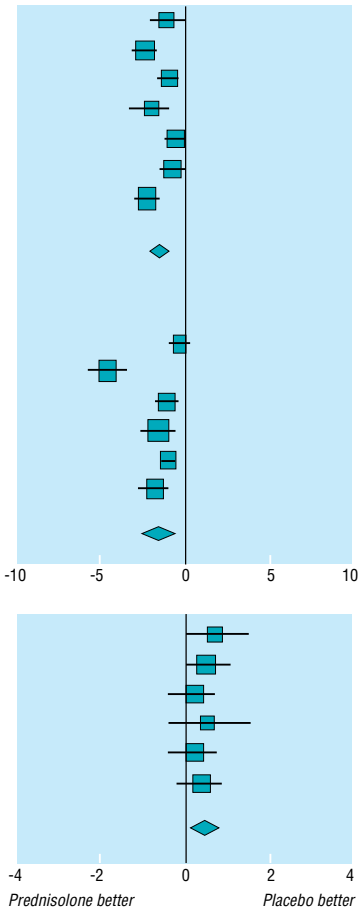


Fig 1 Results of meta-analysis of low dose prednisolone versus placebo for control of rheumatoid arthritis, according to joint tenderness, pain, and grip strength. *If prednisolone is better than control standardised mean difference is negative for joint tenderness and pain but positive for grip strength. Random effects model was used for joint tenderness and pain, and fixed effects model for grip strength

whereas the control group received bed rest and splints for the inflamed joints.[25] Two trials were too poorly reported to be usable for the meta-analysis,[15 16 26] and one reported only on joint size.[29] Three of these four trials found prednisolone or prednisone to be significantly more effective than placebo; the fourth compared prednisolone and indomethacin and gave no numerical data but just reported that there was "no significant difference in response."[26] The four other excluded trials were long term studies that did not report short term data.[4 5 11 13 14] We contacted the authors of these studies to make sure that no short term data had been recorded without being reported. This was confirmed in two cases[4 11]; we were unable to contact any of the authors of the other two studies or of the study that reported only joint size[29] to ensure that no further variables had been recorded.

Ten studies were included in the meta-analysis (table 1).[6 32–43] Most of the studies were quite old and rather small. In all but one[35 36] the criteria of the American Rheumatism Association for classical or definite rheumatoid arthritis were fulfilled. Age, proportion of women, and duration of disease were reported in only half of the studies but they were typical for studies in rheumatoid arthritis: mean age was 55 years, two thirds were women, and the mean (range) duration of disease was 6 (2.1 to 9.6) years. As expected for patients enrolled in steroid trials the severity of the disease, expressed as number of tender joints or Ritchie's tender joint index, was quite pronounced (see fig 1). Prednisolone was used in six trials and prednisone in four.[6 32 34 40 41 43] As prednisone is equipotent with prednisolone and is a pro-drug of prednisolone we have used "prednisolone" as a general term throughout the paper. The doses were 2.5, 3.0, and 7.5 mg in one study each, 10 mg in three studies, and 15 mg in four. The median length of treatment was 1 week.

The randomisation method was not described in any of the trial reports but details were obtained from the authors for one of the studies in which the treatment allocation seemed to have been adequately concealed.[6 32] These authors also provided short term data from their long term trial. All studies were double blind apart from a single blind study in which the patients seemed to have been blinded.[40 41] Eight of the studies were of a crossover design but only one of them reported having tested for sequence effects.[43] Apart from one study[43] the tender joint count was recorded as

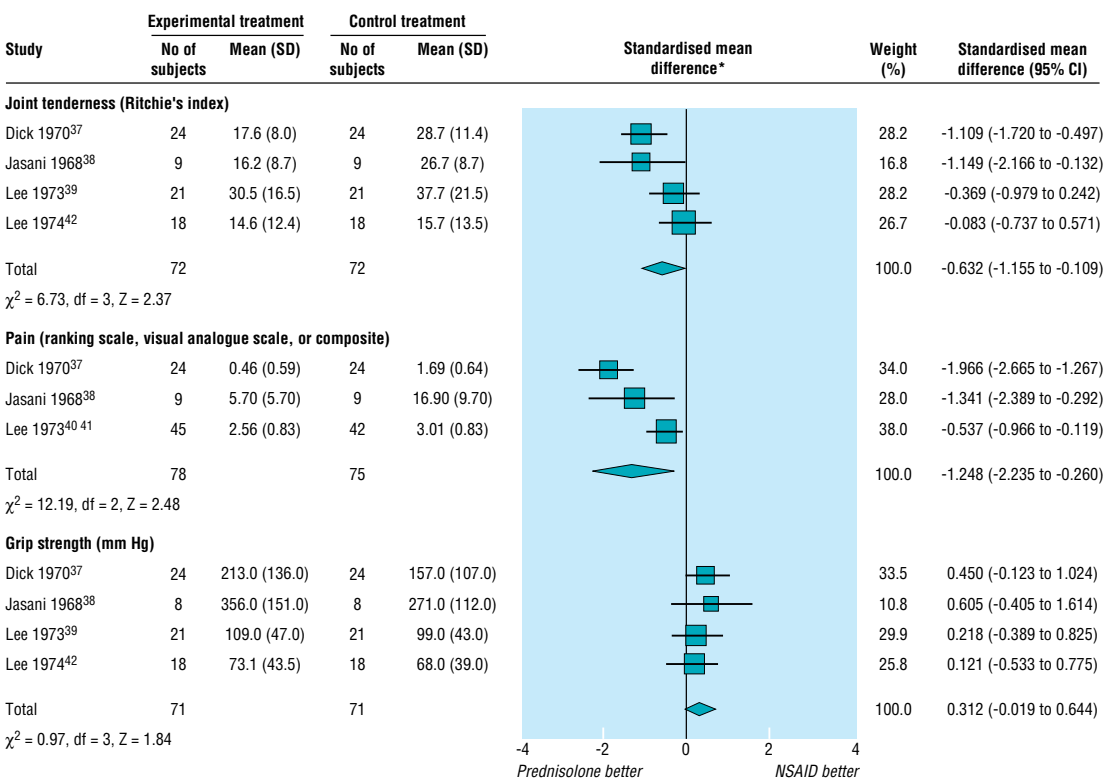


Fig 2 Results of meta-analysis of low dose prednisolone versus non-steroidal anti-inflammatory drugs (NSAIDs) for control of rheumatoid arthritis, according to joint tenderness, pain, and grip strength. *If prednisolone is better than control standardised mean difference is negative for joint tenderness and pain but positive for grip strength. Random effects model was used for joint tenderness and pain, and fixed effects model for grip strength

Ritchie's index; pain was recorded on a ranking scale with 4 or 5 classes in two studies,[35 36 40 41] on a visual analogue scale in two studies,[6 32 33] and as a composite pain index in two studies.[38 43]

The results of the meta-analysis are shown in figures 1 and 2. It should be noted that if prednisolone is better than control, the standardised effect size is negative for joint tenderness and pain but positive for grip strength.

Prednisolone had a clear effect over placebo on joint tenderness (standardised effect size −1.31; 95% confidence interval −1.83 to −0.78), pain (−1.75; −2.64 to −0.87), and grip strength (0.41; 0.13 to 0.69). Measured in the original units, in an analysis of the weighted mean difference the difference between prednisolone and placebo was 12 tender joints (95% confidence interval 6 to 18; test for heterogeneity χ^2 46.42, df=6; P<0.00001). The effect on grip was always measured in mm Hg or in kPa. After conversion of kPa to mm Hg the superiority of prednisolone over placebo was 22 mm Hg (95% confidence interval 5 mm Hg to 40 mm Hg; test for heterogeneity χ^2 5.47, df=5; P=0.36).

Prednisolone also had a greater effect than non-steroidal anti-inflammatory drugs on joint tenderness (−0.63; −1.16 to −0.11), pain (−1.25; −2.24 to −0.26), and grip strength, although the difference in grip strength was not significant (0.31; −0.02 to 0.64). Measured in the original units the difference between prednisolone and non-steroidal anti-inflammatory drugs was 9 tender joints (5 to 12; test for heterogeneity χ^2 4.06, df=3; P=0.26). The effect on grip strength showed a non-significant superiority of prednisolone over non-steroidal anti-inflammatory drugs of 12 mm Hg (−6 mm Hg to 31 mm Hg; test for heterogeneity χ^2 3.03, df=3; P=0.39).

Discussion

Our meta-analysis has shown that low dose prednisolone is not only highly effective but also significantly more effective than non-steroidal anti-inflammatory drugs. The point estimate for the difference in effect between prednisolone and non-steroidal anti-inflammatory drugs on grip strength was 12 mm Hg. It is interesting that the point estimate for the difference in effect between non-steroidal anti-inflammatory drugs and placebo was also found to be 12 mm Hg in an earlier meta-analysis.[44] It was not surprising that the difference in effect on grip strength between prednisolone and non-steroidal anti-inflammatory drugs was not significant as this effect measure is considerably less sensitive to change than pain and joint tenderness.[45]

We used a random effects model for some of the analyses because of heterogeneity. Which model to use is a matter of dispute among statisticians, but the results were not too different if analysed with a fixed effects model, which gave standardised effect sizes for prednisolone versus placebo of −1.23 (−1.51 to −0.95) for joint tenderness and −1.35 (−1.63 to −1.08) for pain, and for prednisolone versus non-steroidal anti-inflammatory drugs of −0.61 (−0.95 to −0.27) for joint tenderness and −0.97 (−1.32 to −0.63) for pain.

Heterogeneity

It is always important to try to explain heterogeneity. Our attempts to do so, however, have been rather unsuccessful. As most of the studies were done more than 20 years ago an obvious reason for the heterogeneity could be that the earlier trials had overestimated the effect—for instance, because of insufficiently concealed randomisation methods.[46] The methodological quality of the trials was acceptable in the whole time span of nearly 30 years, however, and it was, for example, similar to the quality of comparative non-steroidal anti-inflammatory drug trials.[47] In accordance with this there were no time trends for the differences in joint tenderness and pain between prednisolone and placebo. There was marginal heterogeneity ($P = 0.08$) for the difference between prednisolone and non-steroidal anti-inflammatory drugs in joint tenderness, but the heterogeneity disappeared when the analysis was performed in the original units ($P = 0.26$).

Blinding did not seem to have been important for heterogeneity. Only one trial was not double blind, and this trial did not yield larger effect estimates than the other trials. Small trials may exaggerate the effect because of publication bias.[48 49] This possibility could not be studied as the trials were all rather small and contributed similar weights to the meta-analysis. The effect was so pronounced, however, that it would have been unreasonable to plan large trials; in this respect steroid trials resemble trials of non-steroidal anti-inflammatory drugs that have also shown convincingly their superiority over placebo in small crossover trials.[45] One would need to postulate that an unrealistically large number of unpublished trials existed that had shown no effect before the positive effect shown in our meta-analysis would become nullified.

An obvious cause for the heterogeneity could be varying degrees of concomitant treatment with additional non-steroidal anti-inflammatory drugs. Although sometimes stated in trial protocols, it may be difficult to ensure in practice that patients do not take additional drugs. As there was very sparse information on drug intake in the reports this possibility could not be evaluated. Another source could be the use of different measurement scales. Pain, for example, was measured on three different types of scale. They were all ranking scales, and we would therefore definitely have preferred to analyse pain with rank sum tests or as binary data after reduction of the level of measurement. The problem in analysing rank data with parametric methods is not only that they are often far from being normally distributed but also that we do not know the "distances" between the levels on the scale. As the original authors had used parametric statistics we decided to do so as well because our only other option was to discard the data.

Surprisingly, there was no clear relation between dose and effect despite the fact that the doses varied from 2.5 mg to 15 mg daily. It was not the aim of our review, however, to study dose-response relations, which are elucidated more reliably in studies where patients are randomised to different doses. A remarkable effect was seen in a study in which the average dose was only 3 mg daily but where the patients were allowed to start on 7.5 mg when they experienced flares of the arthritis and were advised to take nothing when they were well.[43] This study suggests that it could be an advantage to take steroids intermittently, which would also diminish their adverse effects.

We could be criticised for including crossover trials for which we assumed but could not test that no important carryover effects had occurred. Our arguments for doing this were threefold. Firstly, it is not uncommon in statistical analyses to make necessary assumptions which cannot be properly tested in the data at hand—for example, in multiple regression analyses. Secondly, the problem with crossover trials is not only of a statistical nature, it also has an important ethical dimension. As crossover trials almost without exception are poorly reported and do not allow checks of the assumptions for this design,[47] we would have to discard a vast amount of useful information in the literature in practically all areas of health care if we chose to behave as statistical purists. This would lead to much superfluous research being done, which is not in the best interest of patients or society. Thirdly, and most importantly, one would not expect carryover problems for drugs with relatively quick and reversible symptomatic effects such as steroids or non-steroidal anti-inflammatory drugs in patients with rheumatoid arthritis. In fact in a meta-analysis of non-steroidal anti-inflammatory drugs very similar results were obtained with the two trial designs.[44] For these reasons we believe our approach is justified. Only two studies were of a group comparative design, and the heterogeneity we found could not be explained by type of design.

Included trials

The titles of the included trials were generally quite uninformative and some of the them were not easy to find as they were performed within experiments designed to study other factors. Several of the studies were retrieved from an archive in possession of one of the authors assembled during work on a thesis[50] before the electronic data searches were performed. The authors of the most recent study in this topic[6 32] had found only one of five trials comparing steroids with placebo in long term studies and none of the nine short term trials included in our review. These short term trials were described in 11 reports that were all indexed in Medline with the term for rheumatoid arthritis; in addition, all but one[38] contained the terms for clinical trial or comparative study. Further, all nine trials were identifiable by using the search term "placebo*" *and* ("prednisone" *or* "prednisolone"). This illustrates the value of a systematic and careful search of the literature before starting new clinical trials, and funding bodies and ethical review committees should demand a systematic review of the relevant literature before approving of new clinical research.[51]

Recently, another meta-analysis of low dose corticosteroids (≤15 mg prednisolone daily) in rheumatoid arthritis was published.[52] This meta-analysis looked at moderate term effectiveness and focused on the outcome after 6 months; only two of the included trials were the same as in our meta-analysis.[6 32 43] These authors also noted heterogeneity, but they did not explore possible reasons for it or show the individual results for each trial; they only showed the combined result for each outcome. The weighted mean difference between steroid and placebo was surprisingly small, corresponding to only 2.4 tender joints (four trials,

Table 2 Details of eight trials and two matched cohort studies used in meta-analysis of low dose prednisolone in treatment of rheumatoid arthritis

Study	Equivalent dose of prednisolone	Length of treatment	No of patients taking steroids/ control	Reported major adverse effects (defined by authors)
Randomised trials *v* placebo				
Chamberlain 1976[11]	3 or 5 mg	2 years	30/19	Vertebral fracture in 1 *v* 1; no proved peptic ulcers
Harris 1983[5]	5 mg	6 months	18/16	Two fractures on steroid, no ocular changes; all patients subjected to lumbar spine films and ophthalmic examination
Stenberg 1992[43]	3 mg	3 months	22/22	None (only mild adverse effects, similar to placebo group)
Gestel 1995[6 32]	10 mg	3 months	20/20	No fractures; all patients had lateral spine radiographs taken
Kirwan 1995[4]	7.5 mg	2 years	61/67	None (two cases of hypertension/weight gain on steroid, two with diabetes and hypertension, respectively, on placebo)
Randomised trials *v* aspirin				
Empire Rheumatism Council 1955[13]	15 mg	1 year*	50/50	Hypertension in 2 *v* 0 and indigestion in 1 *v* 5 caused drop out
Joint Committee 1954[20 21]	16 mg†	2 years*	30/32	None (moon face or rubicundity in 11, depression in 5, euphoria in 4 *v* tinnitus in 11, deafness in 10, nausea, dyspepsia or anorexia in 13 reported in first year. Similar adverse effects in second year (one drop out on each drug, no fractures or cataract))
Joint Committee 1959[2]	10 mg‡	2 years*	45/39	Fractures in 2 *v* 1, psychosis in 2 *v* 0, ulcers in 3 *v* 0, infections in 4 *v* 3. All had spinal *x* rays. Several other complications described, most probably unrelated to trial drugs
Matched cohorts				
Saag 1994[53]	⩽15 mg	>12 months	112/112	Survival type analysis; adverse events more common with steroid, see text
McDougall 1994[54]	8 mg	10 years	122/122	Fractures in 31 *v* 19, cataracts in 36 *v* 22, osteonecrosis in 5 *v* 2

*Three year results not analysed because of too many drop outs,[14] treatment not randomised,[22] or too low adherence to randomised treatment.[23]
†Average dose, all started with equivalent of 60 mg prednisolone.
‡Average dose, all started with 20 mg.

95% confidence interval 0.3 to 4.6), while the standardised effect size of 0.90 (−0.18 to 2.00), although not significant, was more comparable to the one we found.

Adverse effects

It is not easy to get a clear picture of the adverse effects of low dose steroids. Five of our short term studies did not report on side effects; one study reported that no side effects occurred[38]; two patients on prednisone had "subjective reactions" in one study[34]; and one patient developed acute psychosis while on prednisone in one study.[40 41] The two remaining studies were moderate term studies from which we extracted short term efficacy data.[6 32 43] These studies did not report short term side effects but are included in the analysis of moderate or long term adverse effects below.

The meta-analysis of moderate term low dose steroid trials did not examine adverse effects at all.[52] The information in the most recently conducted two year placebo controlled trial is also sparse[4]; the aim of this study was to assess the progression of radiological damage, but films were taken only of the hands not of the lumbar spine, which could have detected any compression fractures. We reviewed moderate and long term randomised trials that had compared low dose steroids with placebo or a non-steroidal anti-inflammatory drug. We also identified cohort studies of rheumatoid arthritis that had compared patients treated with steroids with a matched, untreated control group. For this purpose we limited our broad search strategy to *Explode* "glucocorticoids, -synthetic" (adverse-effects) *or Explode* "glucocorticoids" (adverse-effects), combined with *Explode* "arthritis, -rheumatoid" (for all subheadings).

We found eight trials and two matched cohort studies (table 2). Spinal *x* ray photographs were taken of all patients in three of the trials; four fractures were detected in a total of 83 patients randomised to prednisolone and one in 75 patients randomised to placebo. In the five remaining trials, comprising a total of 193 patients taking prednisolone and 190 taking placebo or aspirin, only one fracture with prednisolone and one with placebo were reported. No cases of cataract were reported in the trials. One of the trials was highly atypical as the starting dose was 300 mg cortisone, equivalent to 60 mg prednisolone.[20–22] Its high number of adverse effects may therefore not be representative.

One of the cohort studies used a survival-type analysis and found a large difference in time to first adverse event, with a total of 92 events in the steroid group and 31 in the untreated group.[53] The risk of fracture increased with increasing doses: odds ratio 32.3 (95% confidence interval 4.6 to 220) for >10-15 mg prednisolone daily, 4.5 (2.1 to 9.6) for 5-10 mg, and 1.9 (0.8 to 4.7) for less than 5 mg daily. The overall risks for first event were 3.9 (0.8 to 18.1) for fracture, 8.0 (1.0 to 64.0) for infection, and 3.3 (0.9 to 12.1) for gastrointestinal bleed or ulcer. This study also included patients who received oral steroid "pulses," which do not necessarily lead to the same incidence and severity of adverse effects as continuous low dose treatment. The other cohort study followed two groups of 122 patients for 10 years[54]. Fractures were noted in 31 versus 19 patients, osteonecrosis in 5 versus 2, and cataracts in 36 versus 22 (table 2).

The main problem with studies of matched cohorts is of course that the two groups can never be completely comparable as patients treated with steroids must be expected to be more severely affected than those not treated. This fact may escape notice by traditional measures of morbidity or the difference may be significant for one[54] or more[53] indicators of severity of disease, as in the two cohort studies we reviewed. It is noteworthy, for example, that the first study found a similarly increased risk for fractures as for ulcers,[53] though five meta-analyses of around 100 randomised trials of steroids in various diseases have shown either no increase in risk or, at most, a marginally increased risk of ulcers, which lacks clinical significance.[55] Another meta-analysis of 71 randomised trials, which looked at the risk of infectious complications, showed no increase in risk in patients given less than 10 mg prednisolone daily, and the relative risk for a mean dose under 20 mg was only 1.3 (1.0 to 1.6), which contrasts with the eightfold increased risk in the cohort study.[56] Although the confidence intervals were wide in

Key messages

- Prednisolone in low doses—that is, no more than 15 mg daily—is highly effective in patients with rheumatoid arthritis
- The risk of adverse effects is acceptable in short, moderate, or long term use
- Oral low dose prednisolone may be used intermittently in patients with rheumatoid arthritis, particularly if the disease cannot be controlled by other means
- Further short term placebo controlled trials to study the clinical effect of prednisolone or other oral corticosteroids are no longer necessary

the cohort study, this illustrates the well known dangers of non-randomised comparisons.

Other treatments for rheumatoid arthritis—that is, non-steroidal anti-inflammatory drugs and slow acting antirheumatic drugs—have important adverse effects, which may occasionally even be life threatening. We therefore suggest that short term prednisolone in low doses—that is, not exceeding 15 mg daily—may be used intermittently in patients with rheumatoid arthritis, particularly if they have flares in their disease that cannot be controlled by other means. This suggestion is in accordance with a recent detailed review of the adverse effects of low dose steroids.[57] As prednisolone is highly effective, short term placebo controlled trials to study the clinical effect of low dose prednisolone or other oral corticosteroids are no longer necessary. If additional relevant trials are performed in future—for example, comparison of steroids with non-steroidal anti-inflammatory drugs—they will be included in the electronic version of this meta-analysis,[58] which will be continuously updated.

We are grateful for the unpublished data provided by Anke van Gestel and Roland Laan.

Contributors: PCG wrote the draft meta-analysis protocol and the draft manuscript. HKJ commented on the drafts. Both authors contributed to selection of studies and extraction of data. PCG is guarantor for the study.

Funding: Danish Medical Research Council.

Conflict of interest: None.

1 Hench PS, Kendall EC, Slocumb CH, Polley HF. The effect of a hormone of the adrenal cortex (17-hydroxy-11-dehydrocorticosterone: compound E) and of pituitary adrenocorticotropic hormone on rheumatoid arthritis. *Proc Staff Meet Mayo Clin* 1949;24:181-97.
2 Joint Committee of the Medical Research Council and Nuffield Foundation. A comparison of prednisolone with aspirin or other analgesics in the treatment of rheumatoid arthritis. *Ann Rheum Dis* 1959;18:173-88.
3 Weiss MM. Corticosteroids in rheumatoid arthritis. *Sem Arthritis Rheum* 1989;19:9-21.
4 Kirwan JR and the Arthritis and Rheumatism Council Low-dose Glucocorticoid Study Group. The effect of glucocorticoids on joint destruction in rheumatoid arthritis. *N Engl J Med* 1995;333:142-6.
5 Harris ED, Emkey RD, Nicols JE, Newberg A. Low dose prednisone therapy in rheumatoid arthritis: a double blind study. *J Rheumatol* 1983;10:713-21.
6 Gestel AM van, Laan RFJM, Haagsma CJ, Putte LBA van de, Riel PLCM van. Oral steroids as bridge therapy in rheumatoid arthritis patients starting with parenteral gold. A randomized double-blind placebo-controlled trial. *Br J Rheumatol* 1995;34:347-51.
7 *The Cochrane Library* [database on disk and CD ROM]. Cochrane Collaboration. Oxford: Update Software; 1996-. Updated quarterly.
8 Mulrow CD, Oxman AD. Cochrane Collaboration Handbook, 3rd ed (updated 9 December 1996). Available in *The Cochrane Library* [database on disk and CD ROM]. Cochrane Collaboration; Issue 4. Oxford: Update Software; 1997. Updated quarterly.
9 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clin Trials* 1986;7:177-88.
10 Badia Flores J. Valoracion del GP-40705 en la artritis reumatoide. *Prensa Med Mex* 1969; 34:382-4.
11 Chamberlain MA, Keenan J. The effect of low doses of prednisolone compared with placebo on function and on the hypothalamic pituitary adrenal axis in patients with rheumatoid arthritis. *Rheumatol Rehabil* 1976;15:17-23.
12 Deodhar SD, Dick WC, Hodgkinson R, Buchanan WW. Measurement of clinical response to anti-inflammatory drug therapy in rheumatoid arthritis. *Q J Med* 1973;166:387-401.
13 Empire Rheumatism Council. Multi-centre controlled trial comparing cortisone acetate and acetyl salicylic acid in the long-term treatment of rheumatoid arthritis. *Ann Rheum Dis* 1955;14:353-63.
14 Empire Rheumatism Council. Multi-centre controlled trial comparing cortisone acetate and acetyl salicylic acid in the long-term treatment of rheumatoid arthritis. *Ann Rheum Dis* 1957;16:277-89.
15 Fearnley ME, Masheter HC. A controlled trial of flufenamic acid therapy in rheumatoid arthritis. *Ann Phys Med* 1966;8:204-7.
16 Fearnley ME. An investigation of possible synergism between flufenamic acid and prednisone. *Ann Phys Med* 1966;suppl:109-10.
17 Gum OB. A controlled study of two preparations, paramethasone, propoxyphene, and aspirin and propoxyphene and aspirin in the treatment of arthritis. *Am J Med Sci* 1966;251:328-32.
18 Jick H, Pinals RS, Ullian R, Slone D, Muench H. Dexamethasone and dexamethasone-aspirin in the treatment of chronic rheumatoid arthritis. *Lancet* 1965;ii:1203-5.
19 Jick H, Slone D, Dinan B, Muench H. Evaluation of drug efficacy by a preference technic. *N Engl J Med* 1966;275:1399-403.
20 Joint Committee of the Medical Research Council and Nuffield Foundation. A comparison of cortisone and aspirin in the treatment of early cases of rheumatoid arthritis. *BMJ* 1954;i:1223-7.
21 Joint Committee of the Medical Research Council and Nuffield Foundation. A comparison of cortisone and aspirin in the treatment of early cases of rheumatoid arthritis. *BMJ* 1955;ii:695-700.
22 Joint Committee of the Medical Research Council and Nuffield Foundation. Long-term results in early cases of rheumatoid arthritis treated with either cortisone or aspirin. *BMJ* 1957;i:847-50.
23 Joint Committee of the Medical Research Council and Nuffield Foundation. Comparison of prednisolone with aspirin or other analgesics in the treatment of rheumatoid arthritis. *Ann Rheum Dis* 1960;19:331-7.
24 Kvien TK, Hoyeraal HM, Sandstad B. Assessment methods of disease activity in juvenile rheumatoid arthritis—evaluated in a prednisolone/placebo double-blind study. *J Rheumatol* 1982; 9:696-702.
25 Million R, Kellgren JH, Poole P, Jayson MIV. Long-term study of management of rheumatoid arthritis. *Lancet* 1984;i:812-6.
26 Murthy MHV, Rhymer AR, Wright V. Indomethacin or prednisolone at night in rheumatoid arthritis? *Rheumatol Rehab* 1978;17:8-13.
27 Siegmeth W, Herkner W. Erfahrungen mit einem neuen Kombinationspräparat (Realin) bei der Behandlung akuter rheumatischer Zustandsbilder. *Wien Med Wochenschr* 1974;124:569-73.
28 Slonim RR, Kiem IM, Howell DS, Brown HE. Evaluation of drug responses in the patient severely afflicted with rheumatoid arthritis. *J Florida Med Assoc* 1969;56:336-41.
29 Webb J, Downie WW, Dick WC, Lee P. Evaluation of digital joint circumference measurements in rheumatoid arthritis. *Scand J Rheumatol* 1973;2:127-31.
30 West HF. Rheumatoid arthritis. The relevance of clinical knowledge to research activities. *Abstracts World Med* 1967;41:401-17.
31 Zuckner J, Uddin J, Ramsey RH. Adrenal-pituitary relationships with prolonged low dosage steroid therapy in rheumatoid arthritis. *Missouri Med* 1969;66:649-58.
32 Laan RFJM, Riel PLCM van, Putte LBA van de, Erning LJTO van, Hof MA van't, Lemmens JAM. Low-dose prednisone induces rapid reversible axial bone loss in patients with rheumatoid arthritis. *Ann Intern Med* 1993;119:963-8.
33 Berry H, Huskisson EC. Isotopic indices as a measure of inflammation in rheumatoid arthritis. *Ann Rheum Dis* 1974;33:523-5.
34 Boardman PL, Dudley Hart F. Clinical measurement of the anti-inflammatory effects of salicylates in rheumatoid arthritis. *BMJ* 1967;iv:264-8.
35 Böhm C. Zur medikamentosen Langzeittherapie der primarchronischen Polyarthritis. *Med Welt* 1967;35:2047-50.
36 Schoger GA. Zur beurteilung der Wirkung einer Kombination von Salicylaten und Prednisolon bei rheumatischen Erkrankungen. *Arzneimittel-Forschung* 1968;18:758-60.
37 Dick WC, Nuki G, Whaley K, Deodhar S, Buchanan WW. Some aspects in the quantitation of inflammation in joints of patients suffering from rheumatoid arthritis. *Rheumatol Phys Med* 1970;suppl 10:40-7.
38 Jasani MK, Downie WW, Samuels BM, Buchanan WW. Ibuprofen in rheumatoid arthritis. *Ann Rheum Dis* 1968;27:457-62.
39 Lee P, Jasani MK, Dick WC, Buchanan WW. Evaluation of a functional index in rheumatoid arthritis. *Scand J Rheumatol* 1973;2:71-7.
40 Lee P, Webb J, Anderson J, Buchanan WW. Method for assessing therapeutic potential of anti-inflammatory antirheumatic drugs in rheumatoid arthritis. *BMJ* 1973;ii:685-8.
41 Lee P, Anderson JA, Miller J, Webb J, Buchanan WW. Evaluation of analgesic action and efficacy of antirheumatic drugs. *J Rheumatol* 1976;3:283-94.
42 Lee P, Baxter A, Carson Dick W, Webb J. An assessment of grip strength measurement in rheumatoid arthritis. *Scand J Rheumatol* 1974;3:17-23.
43 Stenberg VI, Fiechtner JJ, Rice JR, Miller DR, Johnson LK. Endocrine control of inflammation: rheumatoid arthritis. Double-blind, crossover clinical trial. *Int J Clin Pharm Res* 1992;12:11-8.
44 Gøtzsche PC. Meta-analysis of grip strength: most common, but superfluous variable in comparative NSAID trials. *Dan Med Bull* 1989;36:493-5.
45 Gøtzsche PC. Sensitivity of effect variables in rheumatoid arthritis: a meta-analysis of 130 placebo controlled NSAID trials. *J Clin Epidemiol* 1990;43:1313-8.

46 Schulz KF, Chalmers I, Hayes RJ, Altman D. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
47 Gøtzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal, antiinflammatory drugs in rheumatoid arthritis. *Controlled Clin Trials* 1989;10:31-56 (erratum:356).
48 Dickersin K, Min Y-I. NIH clinical trials and publication bias. *Online J Curr Clin Trials* 1993 Apr 28;Doc No 50.
49 Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ* 1997;315:640-5.
50 Gøtzsche PC. Bias in double-blind trials (thesis). *Dan Med Bull* 1990;37:329-36.
51 Savulescu J, Chalmers I, Blunt J. Are research ethics committees behaving unethically? Some suggestions for improving performance and accountability. *BMJ* 1996;313:1390-3.
52 Saag KG, Criswell LA, Sems KM, Nettleman MD, Kolluri S. Low-dose corticosteroids in rheumatoid arthritis. A meta-analysis of their moderate-term effectiveness. *Arthritis Rheum* 1996;39:1818-25.
53 Saag KG, Koehnke R, Caldwell JR, Brasington R, Burmeister LF, Zimmerman B, et al. Low dose long-term corticosteroid therapy in rheumatoid arthritis: an analysis of serious adverse events. *Am J Med* 1994;96:115-23.
54 McDougall R, Sibley J, Haga M, Russell A. Outcome in patients with rheumatoid arthritis receiving prednisone compared to matched controls. *J Rheumatol* 1994;21:1207-13.
55 Gøtzsche PC. Steroids and peptic ulcer: an end to the controversy [editorial]? *J Intern Med* 1994;236:599-601.
56 Stuck AE, Minder CE, Frey FJ. Risk of infectious complications in patients taking glucocorticosteroids. *Rev Infect Dis* 1989;11:954-63.
57 Caldwell JR, Furst DE. The efficacy and safety of low-dose corticosteroids for rheumatoid arthritis. *Sem Arthritis Rheum* 1991;21:1-11.
58 Gøtzsche PC, Johansen HK. Short-term low-dose corticosteroids vs placebo and nonsteroidal antiinflammatory drugs in rheumatoid arthritis. In: Tugwell P, Brooks P, Wells G, de Bie R, Bosi-Ferraz M, Gillespie W, eds. Musculoskeletal module. Cochrane Database of Systematic Reviews. *The Cochrane Library*. Cochrane Collaboration; Issue 4. Oxford: Update Software; 1997. Updated quarterly.

(Accepted 18 November 1997)

Effects of the Heartbeat Wales programme over five years on behavioural risks for cardiovascular disease: quasi-experimental comparison of results from Wales and a matched reference area

Chris Tudor-Smith, Don Nutbeam, Laurence Moore, John Catford

Health Promotion Wales, Ffynnon-las, Llanishen, Cardiff CF4 5DZ
Chris Tudor-Smith, *director of research and development*
Department of Public Health and Community Medicine, University of Sydney, NSW 2006, Australia
Don Nutbeam, *professor of public health*
University of Bristol, Department of Social Medicine, Canynge Hall, Bristol BS8 2PR
Laurence Moore, *lecturer in health services research*
Deakin University, Burwood, Victoria 3125, Australia
John Catford, *professor of health strategies*
Correspondence to: Mr Tudor-Smith chris.tudorsmith@hpw.wales.nhs.uk

BMJ 1998;316:818–22

Abstract

Objective: To assess the net 5 year effects of intervention of a community based demonstration project, the Heartbeat Wales programme, on modifiable behavioural risks for prevention of cardiovascular disease.
Design and setting: Quasi-experimental design comparing results from two independent cross sectional population surveys conducted in 1985 and 1990 in Wales and a matched reference area in north east England.
Subjects: Random, stratified samples of people aged 18-64 years (18 538 in 1985 and 13 045 in 1990) in Wales and in north east England (1483 and 4534, respectively).
Intervention: A coordinated range of activities for heart health promotion in Wales entailing public education campaigns along with supportive policy and infrastructure change. In the reference area no additional community heart health promotion was planned, though considerable activity did take place, "contaminating" the reference area.
Main outcome measures: Fifteen self reported behavioural indicators relating to dietary choice, smoking, frequency of exercise, and weight.
Results: Positive changes (for health) in behavioural outcomes were observed among the population in Wales, including a reduction in reported smoking prevalence and improvements in dietary choice. There was no net intervention effect for the programme over and above observed change in the reference area.
Conclusions: No definite conclusions can be drawn concerning the efficacy of the programme in terms of behavioural outcomes. With hindsight, the difficulties of evaluating such a complex multifaceted intervention were underestimated. Further debate on the most appropriate methods for assessing the effectiveness of community based health promotion programmes is called for.

Introduction

Cardiovascular disease remains one of the major causes of morbidity and premature mortality in the United Kingdom.[1] During the 1980s a consensus evolved on the need to reduce this toll of ill health and death through population-wide preventive measures (see, for example, papers by the World Health Organisation[2] and Rose et al[3]). The Welsh Office and the existing national agency for health education, the Health Education Council, agreed to establish a community based demonstration programme in Wales directed towards reducing modifiable behavioural risks for cardiovascular disease.

The programme was publicly launched in 1985 as Heartbeat Wales with three strategic aims: leadership—to coordinate, support, initiate, and monitor action at local and regional levels which would encourage improvements in modifiable behavioural risks for prevention of cardiovascular disease; demonstration—to stimulate, disseminate, and assist the development of strategies and programmes to promote health and prevent cardiovascular disease throughout the United Kingdom; and experimentation—to research, develop, and evaluate a range of new projects and initiatives for heart health promotion and provide feedback on their feasibility and impact.[4]

Heartbeat Wales drew on the experiences of other community based risk reduction programmes for cardiovascular disease, particularly those in Finland and the United States.[5–8] The programme used a range

of established health promotion methods directed towards both changing health behaviours in individuals and achieving environmental, organisational, and policy changes that support healthy choices.[9] Among the resources developed and interventions undertaken by Heartbeat Wales were television series with BBC Wales and HTV such as *Don't Break your Heart*, *Fit for Life*, and the *BBC Diet Programme*; "Quit and Win," a smoking cessation project; food labelling and nutrition education with a major grocery retailer; "Heartbeat Awards," a restaurant and canteen scheme to increase the availability of healthy food choices and smoke free areas; and *Make Health Your Business*, a worksite health promotion programme with CBI Wales.

Further details of the Heartbeat Wales intervention have been published elsewhere.[9–13] To assess behavioural outcomes of the intervention a quasi-experimental evaluation design was adopted on the basis of comparison of change in modifiable behavioural risks for cardiovascular disease in Wales with that in a reference area in the United Kingdom closest in sociodemographic and health profile to Wales at the 1981 census. The reference area selected was north east England (Tyne and Wear, Cleveland, Durham, and North Yorkshire). The Health Education Council indicated that there would be no major additional resources in that area for heart health promotion between 1985 and 1990.

Two population surveys were conducted in 1985 and 1990 in Wales and the reference area. To assist with the interpretation of the findings from these surveys, a range of other studies described elsewhere[14] was also planned for Wales but not the reference area. These studies have suggested that Heartbeat Wales achieved its basic aim of establishing a region-wide approach to the prevention of cardiovascular disease and that many of the key elements of the programme have been taken up and used elsewhere both in the United Kingdom and overseas.[4 10–13 15–18] It has also been shown that there were significant reductions in prevalence of smoking and improvements in food choices between 1985 and 1990 in Wales.[19] This current paper compares these and other changes in modifiable behavioural risks for cardiovascular disease in Wales with those that took place in the reference area over the same time to assess net intervention effects of the Heartbeat Wales programme.

Subjects and methods

Data were collected in random sample, cross sectional surveys during the summer and autumn of 1985 and 1990. In each survey households were selected with a multistage cluster sampling design, within 10 strata defined by the nine Welsh district health authorities and the reference area. Sample size in the 1985 survey was determined ($\alpha = 0.05$, $\beta = 0.2$) to detect a 5% change in prevalence of smoking within each strata by using a two tailed significance test. In 1990 sample size in the reference area was increased to improve the power of analyses that compared Wales with the reference area. Brief interviews were undertaken at each household, and one self completion questionnaire was then left for each resident aged 18-64. Respondents to all three surveys were asked a set of identical questions covering key health related behaviours such as smoking, diet, and physical activity as well as health knowledge and beliefs. In Wales the response rate for the household interview was 88% in 1985 and 79% in 1990 and the self completion response was 67% and 61%, respectively. In the reference area the respective figures were 84% and 77% for the household interview and 64% and 61% for the self completion questionnaire. Altogether, 31 583 questionnaires (18 538 in 1985 and 13 045 in 1990) were returned over the two surveys in Wales, with 6017 (1483 and 4534, respectively) returned in the reference area. Data were weighted before analysis by sex, age group, social class, and population distribution within each strata to minimise bias due to differential response rates between groups. Further details of survey methodology and weighting are available elsewhere.[19]

Table 1 Definition of outcome variables

Indicator	Definition
Health enhancing behaviours	
Chicken	Consume chicken or other poultry ⩾2 days/week
Fish	Consume fish ⩾2 days/week
Fruit	Consume fresh fruit ⩾4 days/week
Green vegetables	Consume green vegetables or salad ⩾4 days/week
Low fat milk	Mainly use skimmed or semi-skimmed milk at home
Wholemeal bread	Mainly use wholemeal bread
Smoking harmful	Smokers who agree that their present level of smoking is harmful to their health
Tried to stop smoking	Smokers who have made a serious attempt to give up in 12 months before survey
Advice on smoking from general practitioner	Daily smokers visiting their general practitioner in 12 months before survey who were advised to cut down or give up
Exercise	Engage in moderate or strenuous activity ⩾2 times/week for >20 minutes each time
Health compromising behaviours	
Butter	Mainly use butter on bread
Fried food	Consume fried food cooked in lard or other solid fat ⩾2 days/week at home
Daily smoking	Smoke daily
Cigarettes/day	Mean No of cigarettes/day smoked by daily smokers
Overweight	Body mass index (kg/m^2) ⩾24 for women; ⩾25 for men

Data analysis

Fifteen indicators were selected as key outcomes for analysis. They represented those health related behaviours that were most consistently targeted during the intervention period and for which measurements were available. These indicators are listed and defined in table 1. Of the 15 indicators, eight represent dietary choices, five are concerned with smoking, and one each with participation in regular exercise and being overweight. Two sets of analyses were undertaken: firstly, at the level of the individual respondent; secondly, at the community level.

Individual level analysis

Standard errors of survey estimates and 95% confidence intervals were estimated on weighted data from the 37 600 completed questionnaires returned in the two surveys by using the SUDAAN (survey data analysis for multistage sample designs) statistical software package.[20] This package uses the Taylor series linearisation method to compute appropriate standard errors for estimates obtained from complex survey designs and takes account of the effects of stratification and clustering on the precision of survey estimates. For each indicator, percentage point changes between

Table 2 Prevalence of key indicators in Wales and reference area, 1985 and 1990, and percentage point changes, 1985-90 (95% confidence intervals). All subjects aged 18-64

	Wales			Reference area		
Key indicator	**1985**	**1990**	**Change**	**1985**	**1990**	**Change**
Chicken	16.5 (15.7 to 17.3)	31.8 (30.5 to 33.1)	15.3 (13.8 to 16.8)	15.8 (13.4 to 18.2)	30.0 (27.7 to 32.3)	14.2 (10.9 to 17.5)
Fish	19.5 (18.7 to 20.2)	27.4 (26.4 to 28.5)	7.9 (6.6 to 9.2)	21.7 (19.4 to 24.0)	28.4 (26.3 to 30.5)	6.7 (3.6 to 9.8)
Fruit	48.3 (47.2 to 49.4)	56.7 (55.4 to 58.0)	8.4 (6.7 to 10.1)	46.2 (41.1 to 51.3)	54.8 (52.5 to 57.0)	8.6 (3.0 to 14.2)
Green vegetables	47.8 (46.5 to 49.1)	55.0 (53.4 to 56.5)	7.2 (5.1 to 9.3)	46.2 (40.0 to 52.5)	55.6 (53.0 to 58.2)	9.4 (2.7 to 16.1)
Low fat milk	16.4 (15.5 to 17.3)	44.1 (42.2 to 46.0)	27.7 (25.7 to 29.7)	20.0 (16.7 to 23.3)*	47.8 (44.9 to 50.7)*	27.8 (23.4 to 32.2)
Wholemeal bread	28.2 (27.0 to 29.4)	35.9 (34.4 to 37.4)	7.7 (5.7 to 9.7)	30.2 (24.9 to 35.4)	37.6 (34.6 to 40.7)	7.4 (1.3 to 13.5)
Smoking harmful	56.4 (55.1 to 57.7)	65.6 (63.9 to 67.3)	9.2 (7.0 to 11.4)	52.9 (48.0 to 57.8)	65.1 (62.3 to 68.0)	12.2 (6.5 to 17.9)
Tried to stop smoking	33.1 (31.7 to 34.5)	34.8 (32.9 to 36.7)	1.7 (–0.6 to 4.0)	35.6 (31.2 to 40.1)	34.5 (31.9 to 37.1)	–1.1 (–6.3 to 4.1)
GP advice on smoking	38.4 (36.5 to 40.2)	49.6 (47.1 to 52.1)	11.2 (8.1 to 14.3)	41.8 (35.4 to 48.2)	49.0 (45.3 to 52.6)	7.2 (–0.2 to 14.6)
Exercise	31.7 (30.9 to 32.5)	33.8 (32.8 to 34.8)	2.1 (0.8 to 3.4)	33.7 (30.3 to 37.2)	36.9 (35.0 to 38.7)†	3.2 (–0.7 to 7.1)
Butter	43.1 (42.0 to 44.2)	26.1 (24.8 to 27.4)	–17.0 (–15.3 to –18.7)	36.8 (34.4 to 39.2)†	22.1 (20.2 to 24.0)†	–14.7 (–11.6 to –17.8)
Fried food	32.5 (31.1 to 34.0)	13.8 (12.6 to 15.0)	–18.7 (–16.8 to –20.6)	42.8 (36.5 to 49.1)†	21.3 (18.6 to 24.0)†	–21.5 (–14.6 to –28.4)
Daily smoking	32.5 (31.4 to 33.6)	27.6 (26.2 to 29.0)	–4.9 (–3.2 to –6.6)	36.0 (30.5 to 41.4)	30.6 (28.3 to 32.9)*	–5.4 (–11.3 to 0.5)
Cigarettes/day‡	17.16 (16.9 to 17.5)	17.14 (16.8 to 17.5)	–0.02 (–0.48 to 0.44)	17.55 (16.8 to 18.3)	17.39 (16.7 to 18.0)	–0.16 (–1.18 to 0.86)
Overweight	43.2 (42.3 to 44.1)	45.7 (44.5 to 46.9)	2.5 (1.0 to 4.0)	39.8 (36.3 to 43.3)	40.9 (39.0 to 42.8)†	1.1 (–2.9 to 5.1)

GP = general practitioner.
*P<0.05 in Z tests for difference in proportions, Wales *v* reference area.
†P<0.01 in Z tests for difference in proportions, Wales *v* reference area.
‡*t* test for difference in means.

1985 and 1990 in Wales and the reference area (with 95% confidence intervals) were calculated.

Community level analysis

Analysis to compare change in Wales with change in the reference area was undertaken at community level by using the nine district health authorities in Wales and the four counties in the reference area as the units for analysis.

The intervention effect to be estimated was defined as the ratio of percentage change between Wales and the reference area $(P_{C90W}/P_{C85W})/(P_{C90R}/P_{C85R})$, where P_{CY} = proportion reporting characteristic C in year (Y = 1985 or 1990) in Wales (W) or the reference area (R).

For each of the 13 analysis units (nine district health authorities in Wales and four counties in the reference area) the prevalence of each of the 14 binary categorical variables was calculated for the baseline and follow up (1990) surveys. Similarly, mean number of cigarettes per day for daily smokers was calculated for each unit in the two surveys. The logarithm of the intervention effect ratio was then estimated by fitting the following model:

$$\ln(P_{C90}) = a + b\cdot Int + c\cdot Z + d\cdot \ln(P_{C85})$$

where P_{CY} = proportion reporting characteristic C in year (Y = 1985 or 1990) in each unit (or mean number of cigarettes per day for daily smokers); Int = exposure to intervention: 1 in Wales, 0 in the reference area; and $Z = \ln(R_{C90}) - \ln(R_{85})$, where R is the proportion of respondents in each unit living in a household where the head of household is in a non-manual occupation. The variable Z was included as a covariate to adjust the analysis for variations between the surveys in the composition of social group of the samples within each unit. The parameter d was included to control for the possibility that the degree of change may be dependent on the baseline value.

Weighted least squares linear regression models were fitted for each of the 15 variables weighted by the mean sample size in each unit over the two surveys. Analyses were undertaken for all respondents, and additional analyses were also run for seven subgroups: men, women, three age groups (18-34, 35-49, 50-64 years), and people living in households where the head of household was in a manual or non-manual occupation. Two tailed *t* tests of the null hypothesis that the parameter b (the logarithm of the intervention effect) was equal to 0 were undertaken and the adjusted intervention effect ratio (exp(b)) and its 95% confidence interval were calculated.

Results

Table 2 shows the prevalence of the 15 key outcome indicators in both Wales and the reference area in 1985 and 1990. These data indicate that in Wales there were positive (for health) changes in all 15 indicators except the proportion overweight, with all changes being significant except for the proportion of smokers who had tried to stop and mean daily cigarette consumption. Similarly, in the reference area there were positive changes in 13 of the 15 indicators, although only nine of these were significant.

Further analysis showed that the baseline prevalence of two of the indicators was significantly lower in Wales than in the reference area—namely, the preference for low fat milk and the consumption of fried food. By 1990 four indicators were significantly lower in Wales than in the reference area; these were preference for low fat milk, consumption of fried food, daily smoking, and participation in exercise. The preference for butter was significantly higher in Wales than in the reference area in both 1985 and 1990, while the proportion overweight was significantly higher in Wales in 1990 only.

Table 3 presents the findings from the community level analyses and indicates that there were two outcome indicators (consumption of fried food and daily smoking) for which there was a consistent intervention effect across all seven subgroups in favour of Wales. There was also one consistent difference (consumption of green vegetables) in favour of the reference area. These effects were each significant in no

more than one subgroup, however, and when the community level analysis was undertaken with all the respondents no significant differences were found.

Discussion

The results indicate important changes in modifiable risks to health among the population in Wales and in the reference area in the north east of England for the period 1985-90. These changes should lead to subsequent reductions in premature death from cardiovascular disease in Wales and the north east of England over the coming decade. As welcome as improvements in smoking levels, dietary habits, and exercise patterns may be, the results do not show clear and consistent net intervention effects of the Heartbeat Wales programme after 5 years in comparison with activities in the reference area.

Sample size and contamination

Interpretation of the results reported here, however, requires a clear understanding of the strengths and weaknesses of the study design and of the context of the intervention. The critical assumption made in the study design was that the contrast between the intervention in Wales and existing activity in the reference area would be large enough and sustained over a 5 year period to show a clear net intervention effect. This was not the case for two reasons. Firstly, the sample size at the baseline measurement in the reference area was too small to give sufficient statistical power for the detection of a likely net intervention effect. Secondly, a previous paper by the authors has documented the diffusion of Heartbeat Wales projects and programmes to the reference area far faster and to a far greater extent than had initially been expected, along with the introduction of additional resources for heart health promotion through the development of the Look After Your Heart project, which was launched across the whole of England in 1987, and the Heartbeat Yorkshire programme, which was conducted in the reference area from 1988.[4] This paper clearly shows the "contamination" of the reference area with initiatives that can be traced back to Heartbeat Wales in whole or part, including the uptake of policy changes in the health services, adoption of several unique projects, and receipt of a number of mass media interventions, especially those developed with BBC Wales. In addition, the paper provides evidence of increases in funding for heart health promotion in the reference area commensurate with increases in Wales during the Heartbeat Wales intervention period.

In retrospect, it was naive to believe that a high profile programme such as Heartbeat Wales could remain in quarantine for such a long period. A direct result of its success as a national demonstration programme was the attenuation of differential exposure to heart health promotion between Wales and the reference area and thus a dilution of any measurable intervention effect. In addition to these identifiable confounding factors are the favourable secular trends in smoking and dietary choices in the United Kingdom as a whole, which further confuse the interpretation of results.[21 22]

Conclusions

Because of these design problems, no definite conclusions can be drawn concerning the efficacy of the Heartbeat Wales programme in terms of behavioural outcomes. Indeed, two directly conflicting conclusions could be drawn, both of which would be compatible with but not proved by the results presented here: on the one hand, the improvements in risk behaviours for cardiovascular disease in Wales suggest that the Heartbeat Wales programme has been effective, with positive changes in the reference area also associated with increased community heart health promotion; on the other hand, the lack of any net intervention effect compared with the reference area

Table 3 Community level regression models: estimates of intervention effect ratio (95% confidence intervals) for seven subgroups and all subjects

	Age (years)			Occupation				
Key indicator	18-34	35-49	50-64	Manual	Non-manual	Men	Women	All subjects
Chicken	1.10 (0.90 to 1.35)	0.92 (0.72 to 1.18)	1.09 (0.83 to 1.42)	1.08 (0.85 to 1.36)	0.87 (0.70 to 1.09)	1.04 (0.85 to 1.28)	0.96 (0.79 to 1.17)	1.01 (0.84 to 1.20)
Fish	0.96 (0.72 to 1.26)	1.16 (0.66 to 2.05)	1.01 (0.82 to 1.24)	1.05 (0.90 to 1.22)	0.94 (0.76 to 1.16)	1.01 (0.80 to 1.27)	1.02 (0.85 to 1.21)	1.00 (0.84 to 1.19)
Fruit	1.04 (0.90 to 1.19)	0.93 (0.80 to 1.07)	0.88 (0.73 to 1.08)	0.99 (0.84 to 1.17)	0.94 (0.82 to 1.08)	0.97 (0.86 to 1.09)	0.98 (0.84 to 1.15)	0.97 (0.86 to 1.09)
Green vegetables	0.96 (0.88 to 1.03)	0.95 (0.82 to 1.10)	0.99 (0.87 to 1.12)	0.96 (0.89 to 1.04)	0.98 (0.90 to 1.06)	0.95 (0.88 to 1.02)	0.97 (0.88 to 1.07)	0.96 (0.89 to 1.03)
Low fat milk	0.98 (0.85 to 1.13)	1.02 (0.91 to 1.16)	1.05 (0.86 to 1.27)	1.03 (0.87 to 1.21)	0.99 (0.81 to 1.20)	1.05 (0.92 to 1.20)	1.03 (0.95 to 1.12)	1.03 (0.94 to 1.14)
Wholemeal bread	0.91 (0.76 to 1.09)	0.91 (0.78 to 1.06)	1.03 (0.87 to 1.21)	0.96 (0.79 to 1.16)	0.93 (0.82 to 1.05)	0.96 (0.76 to 1.20)	0.99 (0.90 to 1.09)	0.98 (0.85 to 1.13)
Smoking harmful	1.02 (0.89 to 1.17)	0.90 (0.69 to 1.17)	0.96 (0.80 to 1.16)	0.97 (0.84 to 1.12)	1.01 (0.78 to 1.30)	0.99 (0.87 to 1.13)	0.94 (0.81 to 1.08)	0.95 (0.85 to 1.07)
Tried to stop smoking	1.06 (0.85 to 1.31)	0.94 (0.76 to 1.16)	1.15 (0.92 to 1.45)	0.95 (0.83 to 1.09)	1.24 (0.94 to 1.63)	1.19 (1.05 to 1.35)*	0.96 (0.81 to 1.13)	1.09 (0.89 to 1.32)
GP advice on smoking	0.97 (0.70 to 1.36)	0.93 (0.74 to 1.18)	1.08 (0.83 to 1.40)	1.03 (0.88 to 1.19)	1.07 (0.78 to 1.46)	1.05 (0.84 to 1.32)	1.02 (0.83 to 1.25)	1.02 (0.95 to 1.09)
Exercise	1.00 (0.90 to 1.11)	1.41 (1.00 to 1.99)	0.94 (0.70 to 1.25)	0.97 (0.87 to 1.07)	0.93 (0.85 to 1.01)	1.00 (0.93 to 1.08)	0.87 (0.74 to 1.01)	0.96 (0.78 to 1.17)
Butter	0.82 (0.60 to 1.13)	1.18 (0.92 to 1.51)	0.96 (0.72 to 1.19)	0.87 (0.73 to 1.03)	1.12 (0.89 to 1.39)	0.96 (0.78 to 1.18)	0.95 (0.83 to 1.09)	0.96 (0.83 to 1.11)
Fried food	0.99 (0.72 to 1.37)	0.86 (0.65 to 1.13)	0.55 (0.37 to 0.83)*	0.78 (0.57 to 1.08)	0.92 (0.63 to 1.34)	0.86 (0.65 to 1.14)	0.84 (0.61 to 1.16)	0.86 (0.64 to 1.14)
Daily smoking	0.93 (0.67 to 1.28)	0.90 (0.73 to 1.10)	0.83 (0.60 to 1.15)	0.91 (0.75 to 1.09)	0.99 (0.82 to 1.19)	0.92 (0.76 to 1.12)	0.91 (0.70 to 1.17)	0.90 (0.73 to 1.10)
Cigarettes/ day	1.01 (0.90 to 1.13)	0.97 (0.85 to 1.10)	0.98 (0.87 to 1.09)	0.99 (0.91 to 1.07)	1.02 (0.93 to 1.13)	1.01 (0.92 to 1.10)	0.96 (0.86 to 1.07)	0.99 (0.93 to 1.07)
Overweight	1.02 (0.77 to 1.36)	1.14 (0.60 to 2.18)	1.04 (0.94 to 1.15)	1.07 (0.89 to 1.28)	1.02 (0.87 to 1.19)	1.06 (0.92 to 1.23)	1.00 (0.86 to 1.15)	1.02 (0.98 to 1.06)

GP=general practitioner.
*P<0.05 in *t* test of log (adjusted intervention effect ratio).

Key messages

- Heartbeat Wales was set up in 1985 as a community based programme to demonstrate risk reduction for cardiovascular disease
- Important changes were observed in modifiable risks for cardiovascular disease in Wales between 1985 and 1990
- There was an unexpectedly rapid uptake of heart health promotion activity in the reference area
- No net intervention effects were found for the programme over and above changes in the reference area
- Improvements in methods of evaluation for community based health promotion programmes are required

suggests that the Heartbeat Wales programme has had no impact additional to secular trends.

These results from the United Kingdom can be set alongside results from other comparable programmes in the United States operating during the same period. Although encouraging results have been reported from the Stanford five city project, these were obtained from a cohort of subjects in the intervention communities. As in Wales, no differences were found in the comparison between independent samples in the intervention population and the reference communities.[6] Similarly, analysis of results from independent samples in the Minnesota and Pawtucket heart health programmes showed that intervention effects were "modest in size and duration and generally within chance levels"[7] and "very limited."[8] The strengths of positive secular trends relating to behavioural risks for cardiovascular disease were cited as a reason why net effects were difficult to identify despite positive results in the intervention communities.[23] Unlike Heartbeat Wales, there is no indication that these studies included active monitoring of the diffusion of the programme in the reference area(s) or the extent of contamination through other unpredicted events.

The major conclusion to be drawn from this study is that the basic quasi-experimental design was inappropriate and insufficiently sensitive to answer the complex research question being asked. By their very nature, successful long term community based programmes can result in complex and wide ranging effects, many of which may be unexpected and not confined to any one predetermined intervention community, making the measurement of any impact and attribution of causality highly problematic.[23–28] Solving these problems will remain a continuing dilemma for advocates of prevention and should be a cause for reflection among academics and researchers concerning appropriate methods for assessing the results from such programmes. New evaluation techniques need to be developed that combine the strengths of quantitative and qualitative research methods and make better use of more proximal outcomes.[29]

We thank Geoffrey Berry, Jim Dwyer, Les Irwig, Russell Luepker, Lucie Richard, Judy Simpson, Chris Roberts, and colleagues at Health Promotion Wales for their help and constructive comments.

CT-S managed the 1990 data collection and coordinated the writing of the paper. DN had overall responsibility for the implementation of the research strategy and contributed to the intervention design and implementation. LM was responsible for and undertook the statistical analyses. JC led the development and implementation of the programme. All authors contributed to the writing of the paper. CT-S and DN are the guarantors for the study.

Funding: Welsh Office.

Conflict of interest: None.

1 Boaz A, Rayner M. *Coronary heart disease statistics*. London: British Heart Foundation, 1995.
2 WHO Expert Committee. Prevention of coronary heart disease. *World Health Organ Tech Rep Ser* 1982;678.
3 Rose G, Ball K, Catford J, James P, Lambert D, Maryon-Davis A, et al. *Coronary heart disease: plans for action*. London: Pitman, 1984.
4 Nutbeam D, Smith C, Murphy S, Catford J. Maintaining evaluation designs in long term community-based health promotion programmes: Heartbeat Wales case study. *J Epidemiol Community Health* 1993;47:127-33.
5 Puska P, Tuomilehto J, Nissinen A, Vartiainen E, eds. *The North Karelia project. 20 years results and experiences*. Helsinki: National Public Health Institute (KTL), 1995.
6 Farquhar JW, Fortmann SP, Flora JA, Taylor CB, Haskell WL, Williams PT, et al. Effects of community wide education of cardiovascular risk factors: the Stanford five-city project. *JAMA* 1990;264:359-65.
7 Luepker RV, Murray DM, Jacobs DR, Mittelmark MB, Bracht N, Carlaw R, et al. Community education for cardiovascular disease prevention: risk factor changes in the Minnesota heart health program. *JAMA* 1994;84:1383-93.
8 Carleton RA, Lasater TM, Assaf AR, Feldman HA, McKinlay S, and the Pawtucket Heart Health Program Writing Group. The Pawtucket heart health program: community changes in cardiovascular risk factors and projected disease risk. *Am J Public Health* 1995;85:777-85.
9 Catford J, Nutbeam D. Heartbeat Wales. In: Williams K, ed. *The community prevention of coronary heart disease*. London: HMSO, 1992:164-75.
10 Clarkson J, Nutbeam D. Introducing healthy catering practice into hospitals: a case study from Wales. *J Nutr Health* 1991;7:101-10.
11 Phillips CJ, Prowle M. Economics of a reduction in smoking: case study from Heartbeat Wales. *J Epidemiol Community Health* 1993;47:215-23.
12 Roberts C, Smith C, Catford J. Quit and win Wales: an evaluation of the 1990 pilot contest. *Tobacco Control* 1993;2:114-9.
13 Murphy S, Powell C, Smith C. A formative evaluation of the Welsh heartbeat award scheme. *J Nutr Health* 1994;9:317-27.
14 Nutbeam D, Catford J. The Welsh heart programme evaluation strategy: progress, plans and possibilities. *Health Promotion* 1987;2:5-18.
15 Nutbeam D, Catford J. Modifiable risks for cardiovascular disease among general practitioners in Wales. *Public Health* 1990;104:353-61.
16 Smith C, Nutbeam D. Assessing non-response bias: a case study from the 1985 Welsh heart health survey. *Health Educ Res* 1990;5:381-6.
17 Pullen E, Nutbeam D, Moore L. Demographic characteristics and health behaviours of consenters to medical examination. Results from the Welsh heart health survey. *J Epidemiol Community Health* 1992;46:455-9.
18 Roberts RJ. Can self-reported data accurately describe the prevalence of overweight? *Public Health* 1995;109:275-84.
19 Smith C, Moore L, Roberts C, Catford J. Health-related behaviours in Wales, 1985-1990. *Health Trends* 1994;26:18-21.
20 Shah BV, Barnwell BG, Bieler GS. *SUDAAN users manual: software for analysis of correlated data, release 6.40*. North Carolina: Research Triangle Institute, 1995.
21 Bennett N, Jarvis L, Rowlands O, Singleton N, Haselden L. *Living in Britain. Results from the 1994 General Household Survey*. London: HMSO, 1996.
22 Ministry of Agriculture, Fisheries and Foods. *Household food consumption and expenditure*. London: HMSO, 1970-1994.
23 Susser M. The tribulations of trials—intervention in communities [editorial]. *Am J Public Health* 1995;85:156-8.
24 Chapman S. Unravelling gossamer with boxing gloves: problems in explaining the decline in smoking. *BMJ* 1993;307:429-32.
25 Green LW, Richard L. The need to combine health education and health promotion: the case of cardiovascular disease prevention. *Promotion and Education* 1993;0:11-7.
26 Winkleby MA. The future of community-based cardiovascular disease intervention studies. *Am J Public Health* 1994;84:1369-72.
27 Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312:1215-8.
28 Speller V, Learmonth A, Harrison D. The search for evidence of effective health promotion. *BMJ* 1997;315:361-3.
29 Nutbeam D. Health outcomes and health promotion: defining success in health promotion. *Health Promotion J Aust* 1996;6:58-60.

(Accepted 18 November 1997)

Correction

Case-control study of risk of cerebral sinus thrombosis in oral contraceptive users who are carriers of hereditary prothrombotic conditions

An error occurred in the title and abstract of this paper by S F T M de Bruijn and others (21 February, pp 589-92). The title should have been: Case-control study of risk of cerebral sinus thrombosis in oral contraceptive users and in carriers of hereditary prothrombotic conditions. The objective in the abstract was wrong in the same respect.

Risk factors for coronary artery disease in non-insulin dependent diabetes mellitus: United Kingdom prospective diabetes study (UKPDS: 23)

R C Turner, H Millns, H A W Neil, I M Stratton, S E Manley, D R Matthews, R R Holman for the United Kingdom Prospective Diabetes Study Group

Abstract

Objective: To evaluate baseline risk factors for coronary artery disease in patients with type 2 diabetes mellitus.

Design: A stepwise selection procedure, adjusting for age and sex, was used in 2693 subjects with complete data to determine which risk factors for coronary artery disease should be included in a Cox proportional hazards model.

Subjects: 3055 white patients (mean age 52) with recently diagnosed type 2 diabetes mellitus and without evidence of disease related to atheroma. Median duration of follow up was 7.9 years. 335 patients developed coronary artery disease within 10 years.

Outcome measures: Angina with confirmatory abnormal electrocardiogram; non-fatal and fatal myocardial infarction.

Results: Coronary artery disease was significantly associated with increased concentrations of low density lipoprotein cholesterol, decreased concentrations of high density lipoprotein cholesterol, and increased triglyceride concentration, haemoglobin A_{1c}, systolic blood pressure, fasting plasma glucose concentration, and a history of smoking. The estimated hazard ratios for the upper third relative to the lower third were 2.26 (95% confidence interval 1.70 to 3.00) for low density lipoprotein cholesterol, 0.55 (0.41 to 0.73) for high density lipoprotein cholesterol, 1.52 (1.15 to 2.01) for haemoglobin A_{1c}, and 1.82 (1.34 to 2.47) for systolic blood pressure. The estimated hazard ratio for smokers was 1.41(1.06 to 1.88).

Conclusion: A quintet of potentially modifiable risk factors for coronary artery disease exists in patients with type 2 diabetes mellitus. These risk factors are increased concentrations of low density lipoprotein cholesterol, decreased concentrations of high density lipoprotein cholesterol, raised blood pressure, hyperglycaemia, and smoking.

Introduction

Patients with type 2 diabetes mellitus have a twofold to threefold increased incidence of diseases related to atheroma,[1] and those who present in their 40s and 50s have a twofold increased total mortality.[2] In the United Kingdom the incidence of macrovascular complications in patients with type 2 diabetes mellitus is twice that of microvascular disease.[3] The greater mortality in patients with type 2 diabetes mellitus than in the general population cannot be explained only by the presence of the three classic risk factors for coronary artery disease—that is, smoking, hypertension, and an increased plasma cholesterol concentration.[4]

Previous prospective studies of patients with type 2 diabetes mellitus had comparatively few patients and cardiovascular end points.[5–11] Many of these studies have not measured the concentration of low density lipoprotein cholesterol, potentially the most important lipid fraction.[12 13]

We report a prospective study of white patients with recently diagnosed type 2 diabetes mellitus. After entry to the United Kingdom prospective diabetes study[14] patients were assessed for baseline risk factors after initial treatment by diet for 3 months. The association of coronary artery disease with baseline risk factors has been assessed irrespective of subsequent treatments.

Subjects and methods

Patients

Between 1977 and 1991, 5102 patients aged 25 to 65 years with type 2 diabetes mellitus based on a fasting plasma glucose concentration >6 mmol/l on two occasions were recruited to the study[14]; 4775 (94%) had fasting plasma glucose values >7.0 mmol/l, which is consistent with the American Diabetic Association's definition of diabetes. Of the 7108 patients originally referred for entry to the study, 2006 (28%) were excluded. These were of a similar age and sex and had a similar fasting plasma glucose concentration as those patients included in the study. The main reasons for exclusion were myocardial infarction in the previous year, current angina or heart failure, accelerated hypertension, proliferative or preproliferative retinopathy, renal failure with a plasma creatinine concentration >175 μmol/l, other life threatening disease such as cancer, an illness requiring systemic steroids, an occupation which precluded insulin treatment, language difficulties, or ketonuria >3 mmol/l suggestive of insulin dependent diabetes mellitus.

The study was approved by the ethics committee in each of the 23 centres: Radcliffe Infirmary, Oxford; Royal Infirmary, Aberdeen; General Hospital, Birmingham; St George's Hospital and Hammersmith Hospital, London; City Hospital, Belfast; North Staffordshire Royal Infirmary, Stoke on Trent; Royal Victoria Hospital, Belfast; St Helier Hospital, Carshalton; Whittington Hospital, London; Norfolk and Norwich Hospital; Lister Hospital, Stevenage; Ipswich Hospital; Ninewells Hospital, Dundee; Northampton Hospital; Torbay Hospital; Peterborough General Hospital; Scarborough Hospital; Derbyshire Royal Infirmary; Manchester Royal Infirmary; Hope Hospital, Salford; Leicester General Hospital; Royal Devon and Exeter Hospital. All patients gave their informed consent to take part in the study.

Diabetes Research Laboratories, Nuffield Department of Medicine, University of Oxford, Radcliffe Infirmary, Oxford OX2 6HE

R C Turner, *professor of medicine*
H Millns, *statistician*
H A W Neil, *lecturer in clinical epidemiology*
I M Stratton, *statistician*
S E Manley, *biochemist*
D R Matthews, *consultant physician in diabetes*
R R Holman, *reader in medicine*

Correspondence to: Professor Turner robert.turner@drl.ox.ac.uk

BMJ 1998;316:823–8

The initial treatment by diet for 3 months was completed by 4178 white patients,[14] with a mean loss of 5 kg body weight, but only 867 (16.9%) were able to achieve a near normal fasting plasma glucose concentration of <6 mmol/l. Cardiovascular disease was evident in 381 (7.5%) patients, of whom 58 (15.2%) had previous myocardial infarction, 144 (37.8%) a definite electrocardiographic Q wave abnormality on Minnesota coding, 7 (1.8%) angina, 1 (0.3%) heart failure, 120 (31.5%) intermittent claudication, and 51 (13.4%) a previous stroke or transient ischaemic attack. Only 2693 (70.9%) of the 3797 datasets could be analysed for all variables, as biochemical measurements were not undertaken until 1981, and some patients had no valid data for one or more of the other variables. Sufficient data were available from 3055 (80.5%) patients for the final Cox model analysis, and 2161 (56.9%) patients had ophthalmic photographic data available for the assessment of the effect of retinopathy.

After the initial treatment diet, patients were randomly allocated to different treatments according to the protocol of the United Kingdom prospective diabetes study.[14] This paper does not include any reference to treatment allocations, actual treatment, or diabetes control during the 10 years of follow up.

Table 1 Baseline characteristics in 2693 white patients with no indication of disease related to atheroma. Results are means (SD) or geometric mean (1 SD interval) unless stated otherwise

Variable	Men (n=1564)	Women (n=1129)
Age (years)	52 (9)	53 (9)
Body mass index (kg/m^2)	27.1 (4.7)	29.4 (6.4)
Waist:hip ratio	0.95 (0.06)	0.87 (0.08)
Systolic blood pressure (mm Hg)	133 (18)	139 (20)
Diastolic blood pressure (mm Hg)	82 (10)	83 (10)
No (%) of patients with hypertension	508 (32)	506 (45)
Fasting plasma glucose (mmol/l)	8.3 (2.8)	9.2 (3.0)
Haemoglobin A_{1c} (%)	6.9 (1.7)	7.4 (1.8)
Total cholesterol (mmol/l)	5.2 (1.0)	5.7 (1.2)
Low density lipoprotein cholesterol (mmol/l)	3.3 (0.9)	3.8 (1.1)
High density lipoprotein cholesterol (mmol/l)	1.04 (0.23)	1.10 (0.24)
Triglyceride (mmol/l)	1.5 (0.9, 2.4)	1.6 (1.0, 2.6)
Fasting plasma insulin (mU/l)	11.3 (6.6, 19.4)	13.2 (7.7, 22.7)
Exercise (No (%) of subjects)		
Sedentary	261 (17)	251 (22)
Moderate	496 (32)	440 (39)
Active	692 (44)	428 (38)
Fit	115 (7)	10 (1)
Smoking (No (%) of subjects)		
Never	344 (22)	501 (44)
Ex-smoker	712 (46)	297 (27)
Current	508 (32)	331 (29)

Table 2 Standardised mortality ratios for 5071 patients recently diagnosed with non-insulin dependent diabetes mellitus compared with general population

	Years since randomisation	No of patients	No observed	No expected	Standardised mortality ratio	P value
Men						
	0 to <5	2992	153	162	0.94	0.78
	5 to <10	2267	161	118	1.36	<0.001
	≥10	566	42	26	1.62	0.002
Women						
	0 to <5	2079	69	72	0.96	0.64
	5 to <10	1581	84	55	1.52	<0.001
	≥10	409	28	12	2.42	<0.0001

Follow up, identification, and classification of end points

Patients were seen every three months in the clinics, and any events that were clinically important were noted. To ascertain whether predetermined criteria for the end points were attained two independent doctors received full information on the patients but without details of treatment.[14] Any discrepancies between the two doctors were adjudicated by two independent senior doctors. All end points were coded according to ICD-9 (international classification of disease, 9th revision).[15]

Three aggregate end points were evaluated: coronary artery disease—that is, fatal or non-fatal myocardial infarction or clinical angina—with an abnormal electrocardiogram at rest or after a treadmill test; fatal or non-fatal myocardial infarction; and fatal myocardial infarction.

Baseline risk factors assessment

Height, waist, and hip circumferences were measured, and the smoking status and amount of exercise taken were ascertained by questionnaire.[14] Retinopathy was assessed by modified Wisconsin grading of four colour photographs of each eye taken at 30° to the horizontal.[14] Blood pressure was recorded as the mean of measurements taken 2 and 9 months after diagnosis with electronic sphygmomanometers. Hypertension was defined as systolic blood pressure ≥160 mm Hg or diastolic blood pressure ≥90 mm Hg, or both, or antihypertensive treatment. After the initial treatment diet, patients were fasted overnight and the following concentrations measured: fasting plasma glucose, haemoglobin A_{1c}, low density lipoprotein cholesterol, high density lipoprotein cholesterol, and insulin.[14 16]

Statistical analyses

Data are reported as means (SD), geometric means (1 SD interval), or percentages. Variables for patients included in or excluded from the analyses were compared by *t* tests, χ^2 tests, or Fisher's exact tests.

Standardised mortality ratio in the patients was calculated from the Office of Population Censuses and Surveys death rates for the general population of England and Wales for the same calendar period, with stratification by sex and age in periods of five years.[17 18]

Age was categorised as <50, 50-54, 55-59, or ≥60 years. Continuous variables were grouped into thirds. The effect of potential risk factors on the three aggregate end points was assessed by Cox proportional hazards models,[19] with censoring at 10 years' follow up. The relation of single risk factors with events after adjustment for age and sex was assessed in 2693 patients with all risk factors measured. Multivariate selection of risk factors was done by a stepwise procedure after adjustment for age and sex. Estimated hazard ratios are represented graphically, with 95% confidence intervals estimated for each group by treating the relative risks as floating absolute risks so that the appropriate variability for each group is shown.[20]

Baseline biochemical and blood pressure values were corrected for any effects from regression to the mean by examining repeat values at six months in 497 patients who had remained on the diet treatment alone. The effect of a unit increment of risk factors on coronary artery disease (1 mmol/l in low density lipoprotein cholesterol concentration, 0.1 mmol/l in

Table 3 Relation of potential risk factors to cardiac end points after adjustment for age and sex, in 2693 white patients with non-insulin dependent diabetes mellitus

Variable	Distribution		Coronary artery disease (n=280)				Non-fatal or fatal myocardial infarction (n=192)				Fatal myocardial infarction (n=79)			
	Lower third	Upper third	P value	Estimated hazard ratios for each third*			P value	Estimated hazard ratios for each third*			P value	Estimated hazard ratios for each third		
Body mass index (kg/m^2)	24.8	29.0	0.65				0.083				0.46			
Waist:hip ratio	0.87	0.94	0.89				0.50				0.89			
Systolic blood pressure (mm Hg)	125	142	0.0032	1	1.52	1.72	0.027	1	1.44	1.70	0.011	1	1.14	2.17
Diastolic blood pressure (mm Hg)	79	87	0.025	1	1.08	1.45	0.0061	1	1.17	1.72	<0.0001	1	0.79	2.09
Hypertension			0.018	1	1.34		0.022	1	1.40		0.008	1	1.83	
Fasting plasma glucose (mmol/l)	7.3	9.7	0.016	1	1.31	1.54	0.13				0.017	1	1.83	2.24
Haemoglobin A_{1c} (%)	6.2	7.5	0.0003	1	1.64	1.78	0.01	1	1.47	1.71	0.0099	1	1.09	2.11
Cholesterol (mmol/l)	4.88	5.77	<0.0001	1	1.79	1.93	0.0086	1	1.62	1.67	0.027	1	1.98	2.01
Low density lipoprotein cholesterol (mmol/l)	3.02	3.89	<0.0001	1	1.48	2.29	0.0002	1	1.44	2.11	0.0043	1	1.06	2.25
High density lipoprotein cholesterol (mmol/l)	0.95	1.15	<0.0001	1	0.87	0.51	0.0085	1	0.84	0.57	0.42			
Triglyceride (mmol/l)	1.22	1.87	<0.0001	1	1.63	1.93	0.011	1	1.40	1.72	0.079			
Insulin (mU/l)	9.7	15.6	0.16				0.022	1	1.18	1.63	0.86			
Exercise (sedentary, moderate, active, fit)			0.54				0.044	1	0.73	0.58 0.74†	0.57			
Smoking (never smoked, ex-smoker, current smoker)			0.016	1	1.18	1.55	0.015	1	1.27	1.74‡	0.57			

*Not given for non-significant data. †For the four categories of exercise. ‡For the three categories of smoking.

high density lipoprotein cholesterol concentration, 10 mm Hg in systolic blood pressure, and 1% in haemoglobin A_{1c}) was estimated by fitting each factor as a continuous variable in a stepwise selected Cox model, leaving other risk factors as categorical variables and adjusting for regression to the mean. Statistical analyses were performed using SAS version 6.1.

Results

Table 1 shows the baseline risk factors assessed for the 2693 white patients who had no previous indication of disease related to atheroma and had complete data when studied after the initial treatment diet. The men who were entered into the United Kingdom prospective diabetes study but excluded from this analysis because of previous cardiovascular disease were older (mean age 56 (SD7) years), had higher systolic blood pressure (mean 141(20) mm Hg), were more likely to be hypertensive (46%), and were more likely to be smokers (10% had never smoked, 51% were ex-smokers, and 39% were current smokers) ($P<0.01$ for each). Women who were excluded from the study were also older (mean age 56 (8) years) and had higher systolic blood pressure (mean 145 (20) mm Hg) ($P<0.01$ for each). Baseline concentrations of total cholesterol, low density lipoprotein cholesterol, high density lipoprotein cholesterol, and triglyceride did not differ in subjects according to the presence of disease related to atheroma ($P>0.01$).

Standardised mortality ratio

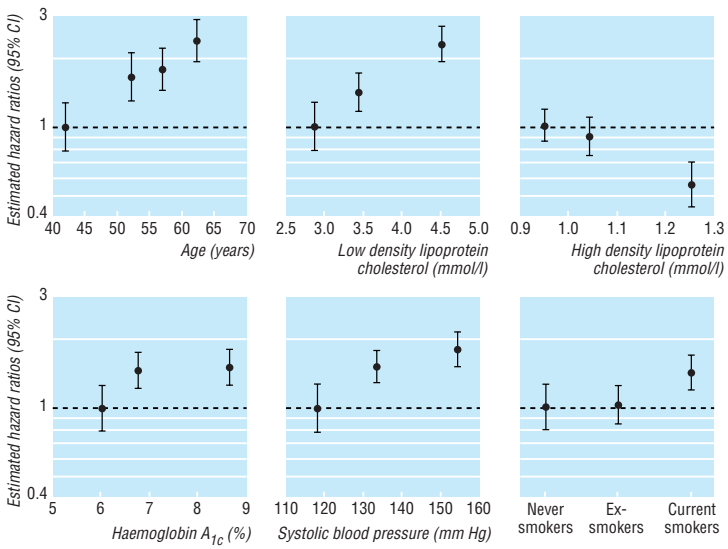
During the first five years of the study the standardised mortality ratio was not very different from that in the general population, possibly because patients with life threatening illnesses were excluded from the study (table 2). After the first five years of the study patients with type 2 diabetes mellitus had an increased total mortality compared with the general population.

Relation of baseline risk factors with adjustment for age and sex

Table 3 shows the relation of potential risk factors, stratified by thirds, to coronary artery disease in 280 patients with end points. Important variables were low density lipoprotein cholesterol concentration, high density lipoprotein cholesterol concentration, and also triglyceride concentration, haemoglobin A_{1c}, systolic blood pressure, fasting plasma glucose concentration, and smoking; each of these had a positive association except high density lipoprotein cholesterol concentration. Similar associations were seen for fatal or non-fatal myocardial infarction (192 patients with end points) and fatal myocardial infarction (79 patients with end points). Retinopathy was associated with fatal

Table 4 Stepwise selection of risk factors, adjusted for age and sex, in 2693 white patients with non-insulin dependent diabetes mellitus with dependent variable as time to first event. P values are significance of risk factor after accounting for all other risk factors in model

Position in model	Coronary artery disease (n=280)		Non-fatal or fatal myocardial infarction (n=192)		Fatal myocardial infarction (n=79)	
	Variable	P value	Variable	P value	Variable	P value
First	Low density lipoprotein cholesterol	<0.0001	Low density lipoprotein cholesterol	0.0022	Diastolic blood pressure	0.0012
Second	High density lipoprotein cholesterol	0.0001	Diastolic blood pressure	0.0074	Low density lipoprotein cholesterol	0.012
Third	Haemoglobin A_{1c}	0.0022	Smoking	0.025	Haemoglobin A_{1c}	0.024
Fourth	Systolic blood pressure	0.0065	High density lipoprotein cholesterol	0.026		
Fifth	Smoking	0.056	Haemoglobin A_{1c}	0.053		



Estimated hazard ratios for significant risk factors for coronary artery disease occurring in 335 out of 3055 diabetic patients (expressed as floating absolute risks)

myocardial infarction (P = 0.005) but not with fatal or non-fatal myocardial infarction (P = 0.124) or with coronary artery diseases (P = 0.082).

Stepwise selection of risk factors

Table 4 shows the selected risk factors with P values from the stepwise multivariate Cox models. Factors that were not important were body mass index, waist to hip ratio, exercise, triglyceride concentration, and fasting plasma glucose or insulin concentration, and these were not included in the final model. The exclusion of triglyceride concentration and not high density lipoprotein cholesterol concentration was not due solely to imprecision of measurements, since baseline triglyceride concentration correlated with the values 6 months later in subjects randomised to, and remaining on, the diet ($r_s = 0.72$ (95% confidence interval 0.68 to 0.76)). This correlation was greater than the correlation between repeated high density lipoprotein measurements ($r_s = 0.52$ (0.45 to 0.58)). Baseline triglyceride concentration correlated with high density lipoprotein cholesterol concentration ($r_s = -0.27$ ($P < 0.0001$)). No significant interaction term was identified between high density lipoprotein cholesterol concentration and low density lipoprotein cholesterol concentration.

Table 5 Estimated hazard ratios (95% confidence intervals) for coronary artery disease in 3055 patients, fitting same explanatory variables for all three dependent variables

Dependent variable	Coronary artery disease	Fatal or non-fatal myocardial infarction	Fatal myocardial infarction
No of patients with event	335	233	103
Age (years)			
<50	1	1	1
50-54	1.65 (1.18 to 2.31)	1.76 (1.15 to 2.70)	5.56 (2.07 to 14.95)
55-59	1.78 (1.29 to 2.46)	2.41 (1.62 to 3.57)	8.70 (3.37 to 22.48)
≥60	2.35 (1.71 to 3.23)	2.80 (1.88 to 4.17)	13.86 (5.43 to 35.41)
Sex			
Women	1	1	1
Men	2.12 (1.65 to 2.73)	2.62 (1.91 to 3.58)	4.13 (2.53 to 6.76)
Low density lipoprotein cholesterol (mmol/l)			
<3.02	1	1	1
≥3.02 to <3.89	1.41 (1.05 to 1.90)	1.41 (1.00 to 2.00)	1.15 (0.67 to 1.97)
≥3.89	2.26 (1.70 to 3.00)	2.11 (1.50 to 2.95)	2.32 (1.41 to 3.81)
High density lipoprotein cholesterol (mmol/l)			
<0.95	1	1	1**
≥0.95 to <1.15	0.90 (0.71 to 1.15)	0.94 (0.70 to 1.26)	1.16 (0.73 to 1.84)
≥1.15	0.55 (0.41 to 0.73)	0.65 (0.47 to 0.91)	0.84 (0.51 to 1.39)
Haemoglobin A_{1c} (%)			
<6.2	1	1	1
≥6.2 to <7.5	1.47 (1.12 to 1.95)	1.26 (0.91 to 1.75)	0.98 (0.58 to 1.65)
≥7.5	1.52 (1.15 to 2.01)	1.42 (1.03 to 1.98)	1.72 (1.06 to 2.77)
Systolic blood pressure (mm Hg)			
<125	1	1*	1*
≥125 to <142	1.52 (1.12 to 2.06)	1.45 (1.01 to 2.08)	1.22 (0.66 to 2.24)
≥142	1.82 (1.34 to 2.47)	1.76 (1.22 to 2.54)	2.36 (1.33 to 4.18)
Smoking			
Never smoked	1	1	1**
Ex-smoker	1.03 (0.77 to 1.37)	1.08 (0.75 to 1.54)	0.65 (0.39 to 1.09)
Current smoker	1.41 (1.06 to 1.88)	1.58 (1.11 to 2.25)	1.03 (0.62 to 1.70)

*Estimated hazard ratios shown for systolic blood pressure despite diastolic blood pressure being included in stepwise selection.
**Variable not included in stepwise selection model.

Estimated hazard ratios

Table 5 and the figure show the Cox model estimated hazard ratios for age and sex and the stepwise selected variables for coronary artery disease in 3055 patients (335 with a coronary artery disease). A similar pattern of hazard ratios was seen for fatal or non-fatal myocardial infarction (233 patients with an event) and fatal myocardial infarction (103 patients with an event). Although diastolic blood pressure had a stronger relation than systolic blood pressure with any myocardial infarction or fatal myocardial infarction, replacement by systolic blood pressure did not affect the results.

Fitting the risk factors for coronary artery disease as continuous variables, with allowance for regression to mean, indicated that for each increment of 1 mmol/l in low density lipoprotein cholesterol concentration there was a 1.57-fold (95% confidence interval 1.37 to 1.79) increased risk of coronary artery disease. For each positive increment of 0.1 mmol/l in high density lipoprotein cholesterol concentration there was a 0.15-fold (0.08 to 0.22) decrease in risk, for each increment of 10 mm Hg in systolic blood pressure a 1.15-fold (1.08 to 1.23) increase, and for each increment of 1% in haemoglobin A_{1c} a 1.11-fold (1.02 to 1.20) increase in risk.

The retinopathy grading was not significantly related to any of the three aggregate end points when added to the multivariate models, including age, sex, and the other risk factors in table 4.

Discussion

This study shows that in patients with type 2 diabetes mellitus increased concentrations of low density lipoprotein cholesterol, decreased concentrations of high density lipoprotein cholesterol, hyperglycaemia, hypertension, and smoking (all measured on completion of a treatment diet after diagnosis), are risk factors for coronary artery disease, defined as fatal and non-fatal myocardial infarction or angina. Previous studies have shown inconsistent results, being dependent on univariate analyses in small studies, with few patients having clinical end points. Total cholesterol concentration was reported to be a risk factor in some[5 7 21] but not other studies,[6 8 22] and most had not measured both low density lipoprotein cholesterol and high density lipoprotein cholesterol concentrations. Hyperglycaemia was similarly reported as a risk factor

in some[5 6 8 11 21 22] but not other studies,[7] and hypertension similarly in some[5 22] but not other studies.[6–8 11 21] The present study shows that patients with type 2 diabetes mellitus have the same risk factors for coronary artery disease as the general population.[23] This study confirms that patients with non-insulin diabetes mellitus have an increased total mortality compared with the general population, although this was not apparent in the initial 5 years, probably because diabetic patients with life threatening illness were excluded from the United Kingdom prospective diabetes study.

Major risk factors

Low density lipoprotein and total cholesterol—An increased concentration of low density lipoprotein cholesterol or total cholesterol at baseline was a major risk factor for coronary artery disease. This is similar to the general population.[13 24 25] Increased concentrations of low density lipoprotein cholesterol may be more pathogenic in patients with type 2 diabetes mellitus than in non-diabetic patients because of the presence of small dense low density lipoprotein cholesterol particles[26] and oxidation of glycated low density lipoprotein cholesterol.[27] The 1.57 increased risk for an increment of 1 mmol/l in low density lipoprotein cholesterol concentration equates to a 36% risk reduction for a decrement of 1 mmol/l , similar to the 31% risk reduction achieved with a 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitor in men with hypercholesterolaemia.[13] The subgroup analysis of the simvastatin study[28] showed that the diabetic patients had similar protection to that of non-diabetic patients.[25] A decreased concentration of high density lipoprotein cholesterol was an independent risk factor for coronary artery disease. The 15% decrease in the risk of coronary artery disease associated with a 0.1 mmol/l increment in high density lipoprotein cholesterol concentration is compatible with the 8-12% reduction reported from prospective American studies.[29] Triglyceride concentration was a risk factor for coronary artery disease after adjustment for age and sex, but it was not an independent risk factor when the other variables were included in the model. This is in accord with other studies, possibly because of the greater biological variability of triglyceride than high density lipoprotein cholesterol measurements.[30] However, we found over 6 months that the concentration of high density lipoprotein cholesterol was more variable than that of triglyceride, possibly because of less precision with the assay (coefficient of variation 6% *v* 2%) and because patients were receiving dietary advice and had a more uniform dietary intake than in the general population. As control of plasma triglyceride and high density lipoprotein cholesterol concentrations is interlinked through lipoprotein lipase and hepatic lipase activities, it may not be feasible to separate the contributions of triglyceride and high density lipoprotein cholesterol to coronary artery disease. Postprandial triglyceride values may have an additional atherogenic role to the fasting values that were measured.[31]

Haemoglobin A_{1c}—There was an increase in risk of coronary artery disease with haemoglobin A_{1c} of $>6.2\%$, the upper range of normal values, in accord with other studies which suggest that glycaemia above the normal range gives an increased risk for macrovascular disease.[32 33] If glycation of proteins was a major pathogenic factor for coronary artery disease the increased risk would be expected to be proportional to the degree of hyperglycaemia.[34] The study showed an increased risk of 11% for each increment of 1% in haemoglobin A_{1c}, similar to the 10% increase in mortality from ischaemic heart disease for an increment of 1% in haemoglobin A_{1c} reported in Wisconsin.[35]

Blood pressure—Increased blood pressure was also a major risk factor for coronary artery disease, with a 15% increased risk for an increase in systolic blood pressure of 10 mm Hg, which was similar to that reported in the general population.[36] Increased blood pressure was a major risk factor for fatal myocardial infarction. This could be because hypertension is a major additional burden to the heart when myocardial infarction ensues. The hypertension in diabetes study in 1148 patients in a factorial design is evaluating whether strict blood pressure control will prevent complications.[37]

Retinopathy—Retinopathy at diagnosis was not a risk factor for cardiovascular disease in a multivariate analysis, although retinopathy was a risk factor for fatal myocardial infarction when only age and sex were adjusted for. Since both retinopathy[38] and microalbuminuria[39] are associated with hyperglycaemia and hypertension, which are also risk factors for coronary artery disease, the previously described association of retinopathy and microalbuminuria with subsequent cardiovascular mortality might reflect the longstanding hypertension and hyperglycaemia that induced both macrovascular and microvascular disease.

Risk factors in type 2 diabetes mellitus

Risk factors for development of coronary artery disease in the general population may not apply once diabetes has developed. Obesity and central obesity,[40] decreased physical activity,[41] and raised insulin concentrations[42] provide an increased risk for cardiovascular disease, but in patients with type 2 diabetes mellitus we found that none of these were major risk factors. These variables are also risk factors for diabetes,[43–45] but this study indicates that once diabetes has developed, hypertension, increased concentrations of low density lipoprotein or decreased concentrations of high density lipoprotein cholesterol and hyperglycaemia measured at baseline are greater risk factors for coronary artery disease than these precipitating factors.

Syndrome X, the association of raised concentrations of glucose, insulin, and triglyceride, decreased concentrations of high density lipoprotein cholesterol, and increased blood pressure, describes a combination of previously reported risk factors for coronary artery disease.[46] In the general population the combination of upper body obesity, glucose intolerance, hypertriglyceridaemia, and hypertension has been termed the deadly quartet.[47] However, a quintet of increased concentrations of low density lipoprotein and decreased concentrations of high density lipoprotein cholesterol, hypertension, hyperglycaemia, and smoking is probably more relevant in patients with type 2 diabetes mellitus.

The cooperation of the patients and many NHS and non-NHS staff at the centres is much appreciated. We thank Professor Eva Kohner, Mr Steve Aldington, Ms Ivy Samuel, and Mrs Caroline Wood for help with the manuscript.

Contributors: RCT and RRH coordinated the study, HM and IMS carried out the statistical analyses, HAWN provided epidemiological advice, DRM coordinated the assessment of

Key messages

- Coronary artery disease is the major cause of mortality in patients with type 2 diabetes mellitus
- Patients without evidence of disease related to atheroma at diagnosis of type 2 diabetes mellitus had an increased standardised mortality ratio compared with the population of the United Kingdom
- 11% of patients in this study had a myocardial infarction or developed angina over a median of 8 years' follow up
- The potentially modifiable risk factors for coronary artery disease were increased concentrations of low density lipoprotein cholesterol, decreased concentrations of high density lipoprotein cholesterol, hypertension, hyperglycaemia, and smoking; these are also risk factors for coronary artery disease in the general population
- Evidence is needed on whether modifying these risk factors will reduce coronary artery disease in patients with type 2 diabetes mellitus

clinical end points, and SEM provided biochemical advice. All authors wrote the paper. RCT will act as guarantor of the study.

Funding: This study received grants from the Medical Research Council; British Diabetic Association; the Department of Health; the National Eye Institute and the National Institute of Digestive, Diabetes and Kidney Disease in the National Institutes of Health, United States; the British Heart Foundation; the Health Promotion Research Trust; Charles Wolfson Charitable Trust; the Alan and Babette Sainsbury Trust; the Clothworkers' Foundation; Oxford University Medical Research Fund Committee; and pharmaceutical companies, including Novo-Nordisk, Bayer, Bristol Myers Squibb, Hoechst, Lilly, Lipha, and Farmitalia Carlo Erba. Additional funding came from Boehringer Mannheim, Becton Dickinson, Owen Mumford, Securicor, Kodak, and Cortecs Diagnostics.

Conflict of interest: None.

1 Garcia MJ, McNamara PM, Gordon T, Kannell WB. Morbidity and mortality in diabetics in the Framingham population. Sixteen year follow-up. *Diabetes* 1974;23:105-11.
2 Panzram G. Mortality and survival in type 2 (non-insulin-dependent) diabetes mellitus. *Diabetologia* 1987;30:123-31.
3 United Kingdom Prospective Diabetes Study Group. UKPDS 17. A nine-year update of a randomized, controlled trial on the effect of improved metabolic control on complications in non-insulin-dependent diabetes mellitus. *Ann Intern Med* 1996;124:136-45.
4 Stamler J, Vaccaro O, Neaton JD, Wentworth D. Diabetes, other risk factors, and 12 year cardiovascular mortality for men screened in the Multiple Risk Factor Intervention Trial. *Diabetes Care* 1993;16:434-44.
5 Gall MA, Borch-Johnsen K, Hougaard P, Nielsen FS, Parving HH. Albuminuria and poor glycaemic control predict mortality in NIDDM. *Diabetes* 1995;44:1303-9.
6 Kuusisto J, Mykkänen L, Pyörälä K, Laakso M. NIDDM and its metabolic control predict coronary heart disease in elderly subjects. *Diabetes* 1994;43:960-7.
7 Fontbonne A, Eschwège E, Cambien F, Richard JL, Ducimetière P, Thibult N, et al. Hypertriglyceridaemia as a risk factor of coronary heart disease mortality in subjects with impaired glucose tolerance or diabetes. Results from the 11-year follow-up of the Paris Prospective Study. *Diabetologia* 1989;32:300-4.
8 Uusitupa MI, Niskanen LK, Siitonen O, Voutilainen E, Pyörälä K. Ten-year cardiovascular mortality in relation to risk factors and abnormalities in lipoprotein composition in type 2 (non-insulin-dependent) diabetic and non-diabetic subjects. *Diabetologia* 1993;36:1175-84.
9 Butler WJ, Ostrander LD, Carman WJ, Lamphiear DE. Mortality from coronary heart disease in the Tecumseh study. Long-term effect of diabetes mellitus, glucose tolerance and other risk factors. *Am J Epidemiol* 1985;121:541-7.
10 Heyden S, Heiss G, Bartel AG, Hames CG. Sex differences in coronary mortality among diabetics in Evans County, Georgia. *J Chronic Dis* 1980;33:265-73.
11 Laakso M, Lehto S, Penttila I, Pyörälä K. Lipids and lipoproteins predicting coronary heart disease mortality and morbidity in patients with non-insulin-dependent diabetes. *Circulation* 1993;88:1421-30.
12 Law MR, Wald NJ, Wu T, Hackshaw A, Bailey A. Systematic underestimation of association between serum cholesterol concentration and ischaemic heart disease in observational studies: data from the BUPA study. *BMJ* 1994;308:363-6.
13 Shepherd J, Cobbe SM, Ford I, Isles CG, Lorimer AR, MacFarlane PW, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. *N Engl J Med* 1995;333:1301-7.
14 United Kingdom Prospective Diabetes Study Group. UKPDS VIII. Study design, progress and performance. *Diabetologia* 1991;34:877-90.
15 World Health Organisation. *International classification of procedures in medicine*. Geneva: World Health Organisation, 1978.
16 United Kingdom Prospective Diabetes Study Group. UKPDS XI. Biochemical risk factors in type 2 diabetic patients at diagnosis compared with age-matched normal subjects. *Diabetic Med* 1994;11:534-4.
17 Population Censuses and Surveys Office. *Mortality statistics: review of the registrar general on deaths in England and Wales*. London: HMSO, 1994.
18 Coleman MP, Herman C, Douglas A. *Person-years (PYRS): a FORTRAN program for cohort study analysis*. Lyons: International Agency for Research on Cancer, 1989.
19 Cox DR. Regression models and life-tables. *J R Stat Soc Series B* 1972;34:187-201.
20 Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Stat Med* 1991;10:1025-35.
21 Lehto S, Ronnemaa T, Haffner SM, Pyörälä K, Kallio V, Laakso M. Dyslipidemia and hyperglycemia predict coronary heart disease events in middle-aged patients with NIDDM. *Diabetes* 1997;48:1354-9.
22 Walters DP, Gatling W, Houston AC, Mullee MA, Julious SA, Hill RD. Mortality in diabetic subjects: an eleven-year follow-up of a community-based population. *Diabetic Med* 1994;11:968-73.
23 Kannel WB, Wilson PW. An update on coronary risk factors. *Med Clin North Am* 1995;79:951-71.
24 Law MR, Wald NJ, Thompson SG. By how much and how quickly does reduction in serum cholesterol concentration lower risk of ischaemic heart disease? *BMJ* 1994;308:367-72.
25 Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 1994;344:1383-9.
26 Austin MA, Breslow JL, Hennekens CH, Buring JE, Willett WC, Krauss PM. Low-density lipoprotein subclass patterns and risk of myocardial infarction. *JAMA* 1988;260:1917-21.
27 Kawamura M, Heinecke JW, Chait A. Pathophysiological concentrations of glucose promote oxidative modification of low density lipoprotein by a superoxide-dependent pathway. *J Clin Invest* 1994;94:771-8.
28 Pyörälä K, Pedersen TR, Kjekshus J. The effect of cholesterol lowering with simvastatin on coronary artery diseases in diabetic patients with coronary heart disease. *Diabetes* 1995;44(suppl):35A.
29 Gordon DJ, Probstfield JL, Garrison RJ, Neaton JD, Castelli WP, Knoke JD, et al. High-density lipoprotein cholesterol and cardiovascular disease. Four prospective American studies. *Circulation* 1989;79:8-15.
30 Schectman G, Sasse E. Variability of lipid measurements: relevance for the clinician. *Clin Chem* 1993;39:1495-1503.
31 Stensvold I, Tverdal A, Urdal P, Graff-Iversen S. Non-fasting serum triglyceride concentration and mortality from coronary heart disease and any cause in middle aged Norwegian women. *BMJ* 1993;307:1318-22.
32 Jarrett RJ, Keen H. Hyperglycaemia and diabetes mellitus. *Lancet* 1976;2:1009-12.
33 Folsom AR, Szklo M, Stevens J, Liao F, Smith R, Eckfeldt J. A prospective study of coronary heart disease in relation to fasting insulin, glucose, and diabetes. *Diabetes Care* 1997;20:935-42.
34 Brownlee M. Glycation products and the pathogenesis of diabetic complications. *Diabetes Care* 1992;15:1835-43.
35 Klein R. Hyperglycemia and microvascular and macrovascular disease in diabetes. *Diabetes Care* 1995;18:258-68.
36 MacMahon S, Peto R, Cutler J, Collins R, Sorlie P, Neaton J, et al. Blood pressure, stroke, and coronary heart disease. Part 1. Prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet* 1990;335:765-74.
37 Hypertension in Diabetes Study Group. HDS 4: therapeutic requirements to maintain tight blood pressure control. *Diabetologia* 1996;39:1554-61.
38 Stratton IM, Kohner E, Aldington S, Matthews DR, Turner RC. Prevalence of diabetic retinopathy at diagnosis of non-insulin dependent diabetes in 2964 white Caucasian patients and association with hypertension, hyperglycaemia and impaired beta-cell function. *Diabetes* 1995;43:S434.
39 United Kingdom Prospective Diabetes Study Group. UKPDS X. Urinary albumin excretion over 3 years in diet-treated type 2 (non-insulin-dependent) diabetic patients, and association with hypertension, hyperglycaemia and hypertriglyceridaemia. *Diabetologia* 1993;36:1021-9.
40 Byers T. Body weight and mortality. *N Engl J Med* 1995;333:723-4.
41 Manson JE, Willett WC, Stampfer MJ, Colditz GA, Hunter DJ, Hankinson SE, et al. Body weight and mortality among women. *N Engl J Med* 1995;333:677-85.
42 Desprès JP, Lamarche B, Mauriège P, Cantin B, Dagenais GR, Moorjani S, et al. Hyperinsulinemia as an independent risk factor for ischemic heart disease. *N Engl J Med* 1996;334:952-7.
43 Modan M, Karasik A, Halkin H, Fuchs Z, Lusky A, Shitrit A, et al. Effect of past and concurrent body mass index on prevalence of glucose intolerance and type 2 (non-insulin-dependent) diabetes and on insulin response. The Israel study of glucose intolerance, obesity and hypertension. *Diabetologia* 1986;29:82-9.
44 Desprès JP, Moorjani S, Lupien PJ, Tremblay A, Nadeau A, Bouchard C. Regional distribution of body fat, plasma lipoproteins, and cardiovascular disease. *Arteriosclerosis* 1990;10:497-511.
45 Manson JE, Rimm EB, Stampfer MJ, Colditz GA, Willett WC, Krolewski AS, et al. Physical activity and incidence of non-insulin-dependent diabetes mellitus in women. *Lancet* 1991;338:774-8.
46 Reaven GM. Role of insulin resistance in human disease. *Diabetes* 1988;37:1595-1607.
47 Kaplan NM. The deadly quartet. Upper-body obesity, glucose intolerance, hypertriglyceridemia, and hypertension. *Arch Intern Med* 1989;149:1514-20.

(Accepted 24 October 1997)

Association between plasma concentrations of plasminogen activator inhibitor-1 and survival in patients with colorectal cancer

Hans Jørgen Nielsen, Helle Pappot, Ib Jarle Christensen, Nils Brünner, Ole Thorlacius-Ussing, Flemming Moesgaard, Keld Danø, Jan Grøndahl-Hansen

Invasion by cancer cells requires proteases such as the serine protease plasmin to degrade the cellular matrix. Plasmin is formed from its zymogen, plasminogen, a reaction catalysed by urokinase type plasminogen activator—which is implicated in invasion[1]—and partly regulated by plasminogen activator inhibitors. The active form of the inhibitor complexes with free and receptor bound active urokinase plasminogen activator and is bound by vitronectin in plasma and extracellular matrix.[2]

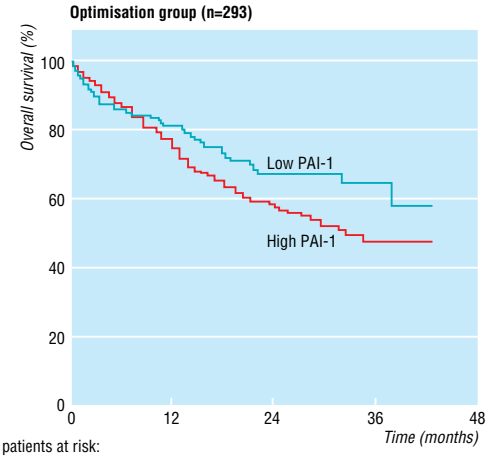
A high concentration of plasminogen activator inhibitor-1 in biopsy specimens from tumours is associated with a poor prognosis,[3] and some patients with ovarian cancer have raised plasma concentrations of plasminogen activator inhibitor-1.[4] We studied the prognostic importance of plasma concentrations of plasminogen activator inhibitor-1 in patients with colorectal cancer.

Subjects, methods, and results

Plasma was collected preoperatively as previously described[5] from 609 patients having elective surgery for colorectal cancer. Plasma concentrations of plasminogen activator inhibitor-1 were measured by a sandwich enzyme linked immunosorbent assay (ELISA) using two monoclonal antibodies.[3] The concentration was expressed as interim units of plasminogen activator inhibitor-1/mg protein.[3]

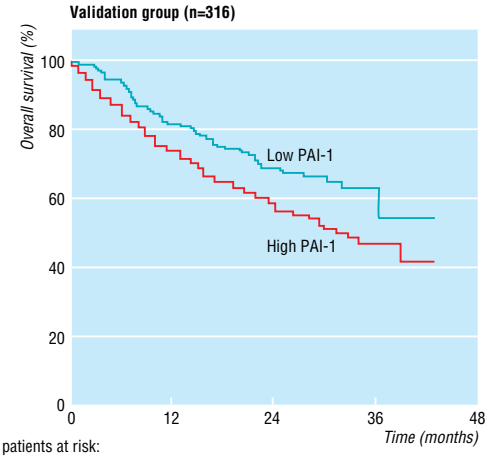
All patients had histologically verified colorectal cancer and complete clinical data. The median follow up time was 25 months (range 13-40). Patients were randomised into two groups. Data on 293 patients (optimisation group) were used to determine the optimal cut off value for plasminogen activator inhibitor-1 in relation to survival using Cox's proportional hazard model, and data on 316 patients (validation group) were used to validate the results obtained from the optimisation group.

High plasma concentrations of plasminogen activator inhibitor-1 were associated with increasing severity of disease (Dukes's stage; χ^2 test, $P = 0.001$). The best cut off value for plasminogen activator inhibitor-1 was 0.5 interim units/mg of protein. With this value the hazard ratio was 1.5 for patients with high concentrations of plasminogen activator inhibitor-1 (178/293 (61%)) compared with those with low concentrations (115/293 (39%)). Applying this value to the validation group gave similar results (hazard ratio 1.5 (95% confidence interval 1.1 to 2.2); $P = 0.02$; 179/316 (57%) *v* 137/316 (43%)) (figure). Cox analysis of the 316 patients in the validation group showed that Dukes's stage was the strongest prognostic variable (hazard ratio 2.9 (2.3 to 3.7)), followed by age (hazard ratio 1.5 (1 to 2.1)).



No of patients at risk:

Low PAI-1	115	93	55	15
High PAI-1	178	133	83	16



No of patients at risk:

Low PAI-1	137	112	70	8
High PAI-1	179	132	81	18

Survival in patients with colorectal cancer according to preoperative plasma concentrations of plasminogen activator inhibitor-1 (PAI-1). See text for definition of groups

Comment

This study shows that high preoperative plasma concentrations of plasminogen activator inhibitor-1 are associated with shorter survival in patients with colorectal cancer. The validity of this result is strongly supported by the fact that the best cut off value for plasminogen activator inhibitor-1 obtained from one patient population gave similar prognostic information about a second independent population. It is further supported by the close correlation between high plasma concentrations of plasminogen activator

Editorial by Verspaget

Department of Surgical Gastroenterology, Hvidovre Hospital, DK-2650, Hvidovre, Denmark
Hans Jørgen Nielsen, *senior researcher*
Flemming Moesgaard, *consultant*
Finsen Laboratory, Rigshospitalet, Strandboulevarden 49, DK-2100, Copenhagen, Denmark
Helle Pappot, *research student*
Ib Jarle Christensen, *associate statistician*
Nils Brünner, *associate professor*
Keld Danø, *professor*
Jan Grøndahl-Hansen, *senior researcher*
Department of Surgical Gastroenterology, Ålborg Hospital, DK-9000, Ålborg, Denmark
Ole Thorlacius-Ussing, *consultant*
Correspondence to: Dr Nielsen

BMJ 1998;316:829–30

inhibitor-1 and increasing severity of disease according to Dukes's stage, which is an established predictor of poor prognosis in patients with colorectal cancer.

We thank the RANX05 Colorectal Cancer Study Group for the collection of the plasma samples.

Contributors: HJN and NB had the original idea for and planned the study. HJN was also responsible for collecting the samples and patient data. FM established the database and participated in planning the clinical trial. JG-H developed the enzyme linked immunosorbent assay and analysed all the samples. IJC and HP were responsible for the statistical analyses of the data. The paper was written jointly by NB, KD, JG-H, OT-U, and HJN. HJN and NB are guarantors of the study.

Funding: This study received financial support from Glaxo Group Research, United Kingdom; the Kornerup Foundation; the Ingeborg Roikjer Foundation; the University of Copenhagen; the Gerda and Aage Haench Foundation; and the Danish Cancer Society.

Conflict of interest: None.

1 Danø K, Andreasen PA, Grøndahl-Hansen J, Kristensen P, Nielsen LS, Skriver L.Plasminogen activators, tissue degradation and cancer. *Adv Cancer Res* 1985;44:139-266.
2 Andreasen PA, Georg B, Lund LR, Riccio A, Stacey SN. Plasminogen activator inhibitors: hormonally regulated serpins. *Mol Cell Endocrinol* 1990;68:1-19.
3 Grøndahl-Hansen J, Christensen IJ, Rosenquist C, Brünner N, Mouridsen HT, Danø K, et al. High level of urokinase-type plasminogen activator and its inhibitor PAI-1 in cytosolic extracts of breast carcinomas are associated with poor prognosis. *Cancer Res* 1993;53:2513-21.
4 Casslen B, Bossmar T, Lecander I, Åstedt B. Plasminogen activators and plasminogen activator inhibitors in blood and tumour fluids of patients with ovarian cancer. *Eur J Cancer* 1994;30A:1302-9.
5 Meijer P, Pollet DE, Wauters J, Kluft C. Specificity of antigen assays of plasminogen activator inhibitor in plasma: Innotest plasminogen activator inhibitor-1 immunoassay evaluated. *Clin Chem* 1994;40:110-5.

(Accepted 10 April 1997)

How much does relapse after one year erode effectiveness of smoking cessation treatments? Long term follow up of randomised trial of nicotine nasal spray

John A Stapleton, Gay Sutherland, Michael A H Russell

Institute of Psychiatry, Tobacco Research Section, National Addiction Centre, London SE5 8AF

John A Stapleton, *senior lecturer*

Gay Sutherland, *lecturer*

Michael A H Russell, *professor*

Correspondence to: Dr Stapleton j.stapleton@iop.bpmf.ac.uk

BMJ 1998;316:830–1

Recent research on treatments to stop smoking has focused almost entirely on nicotine replacement, and several meta-analyses testify to the efficacy of four delivery systems.[1] Although the ultimate goal of treatment is lifelong cessation, few trials have published results of abstinence beyond one year. Little consideration has therefore been given to whether the treatment is effective in reducing the major health risks of smoking. This effect would become evident only after many years of abstinence. Our randomised trial showed that the use of nicotine nasal spray compared with a placebo spray was associated with more than double the number of abstainers at one year.[2] We report the results from a longer term follow up to estimate the impact of relapse after one year on effectiveness.

Subjects, methods, and results

A total of 227 heavy smokers entered the trial; 116 were given the nicotine spray and 111 the placebo. Of these, 47 sustained abstinence from smoking for 1 year. They constituted the long term follow up group; 33 were in the nicotine group, 14 in the placebo group. Criteria for long term sustained abstinence were the same as for the first year. Since the follow up was completed mainly over a 2 month period, the time interval from randomisation varied according to when the smoker entered the trial over 15 months. Standard survival methods were used to analyse the data. Survival times of those who were not contacted beyond 1 year (3 subjects in the nicotine group, 2 in the placebo group) and those who had successfully given up were censored at their last follow up. The Kaplan-Meier method was used to estimate cumulative abstinence up to $3\frac{1}{2}$ years.

Mean follow up period was 3 years 4 months (range $2\frac{1}{2}$ to $4\frac{2}{3}$ years) and was shorter by 21 days for the nicotine group. All observed relapses occurred within $3\frac{1}{2}$ years. The table shows that the nicotine spray maintained an advantage over placebo up to $3\frac{1}{2}$ years. Relapse after 1 year's abstinence was similar in the two groups and totalled 23% at 2 years, 38% at 3 years and 48% (95% confidence interval 32% to 64%) at $3\frac{1}{2}$ years. Although subjects had been recommended to use the spray for three months only, they were allowed to continue for 1 year. Of those remaining abstinent in the nicotine group, 19 used the spray for 1 year and 14 for <1 year (range 1-39 weeks). There was no difference in relapse after 1 year in the nicotine group between those who used the spray for 1 year and those who stopped earlier (difference 5%, 95% confidence interval −33% to 43%).

Results of long term follow up of randomised trial of nicotine nasal spray

	Nicotine spray (n=116)	Placebo spray (n=111)	Difference in % (95% CI)
% (No) who sustained abstinence to 1 year	28.4 (33)	12.6 (14)	15.8 (5.6 to 26.1)
Sustained abstinence to 3.5 years (%)*	15.4	6.1	9.3 (0.88 to 17.4)
Cumulative relapse between 1 and 3.5 years (%)*	45.9	52.1	–6.2 (–41.0 to 28.8)

*Kaplan-Meier estimates.

Comment

Our results show that the spray is an effective aid to long term smoking cessation and that those who used the spray for 1 year had a similar relapse profile to those who stopped using it earlier. They also indicate substantial relapse after the time that most studies have completed their final follow up to assess treatment efficacy. Although the success ratio of active to placebo treatment (about 2.5) was unchanged by relapse, the absolute difference was reduced considerably, and hence the estimated number needed to treat to achieve each success was increased (from 6.3 to 10.8).

Our relapse rate is similar to that in a trial of nicotine patches (37% between years 1 and 3)[3] and in a study using supportive counselling and nicotine gum for 5 years (40% between years 1 and 5).[4] High relapse rates after 1 year are also common in those not attending for treatment. A large general population survey estimated a relapse rate of 35% from non-validated self reports of the duration of abstinence.[5]

Success rates after 1 year or less of follow up substantially overestimate lifelong cessation after a single treatment episode.

Contributors: MAHR, GS, and JAS designed the original study, which GS conducted. JAS and GS designed the long term follow up phase, which GS conducted. JAS analysed the data and wrote the text of the paper. JAS and GS will act as guarantors for the paper.

Funding: Medical Research Council.
Conflict of interest: None.

1 Silagy C, Mant D, Fowler G, Lodge M. Meta-analysis on efficacy of nicotine replacement therapies in smoking cessation. *Lancet* 1994;343:139-42.
2 Sutherland G, Stapleton JA, Russell MAH, Jarvis MJ, Hajek P, Belcher M, et al. Randomised controlled trial of nasal nicotine spray in smoking cessation. *Lancet* 1992;340:324-9.
3 Mikkelsen KL, Tønnesen P, Nørregaard J. Three-year outcome of two-and three-year sustained abstainers from a smoking cessation study with nicotine patches. *J Smok Rel Dis* 1994;5:95-100.
4 Anthonisen NR, Connett JE, Kiley JP, Altose MD, Bailey WC, Buist AS, et al. Effects of smoking intervention and the use of an inhaled anticholinergic bronchodilator on the rate of decline of FEV_1. *JAMA* 1994;272:1497-505.
5 US Department of Health and Human Services. Results from the NHANES-I population survey. In: *The health benefits of smoking cessation: a report of the surgeon general.* Washington, DC: DHHS, 1990. (Publication No CDC 90-8416.)

(Accepted 8 July 1997)

Management of deliberate self poisoning in adults in four teaching hospitals: descriptive study

Navneet Kapur, Allan House, Francis Creed, Eleanor Feldman, Trevor Friedman, Elspeth Guthrie

Deliberate self poisoning accounts for 100 000 hospital admissions in England and Wales every year, and its incidence is increasing.[1] One per cent of patients kill themselves in the year following attendance.[2] Good services to manage deliberate self poisoning in general hospitals might therefore help to achieve the targets set out by the Health of the Nation strategy to reduce suicide rates. Existing services have not been planned coherently; the care provided by hospitals varying greatly, even in the same region.[3] We assessed the management of self poisoning in four teaching hospitals in England by using standardised methods of notification.

Subjects, methods, and results

We prospectively identified all patients over 16 years of age who attended four teaching hospitals in Leeds, Leicester, Manchester, and Nottingham for deliberate self poisoning during 4 weeks (November to December 1996). We obtained data by examining computerised databases on wards and in the accident and emergency department, referral ledgers, accident and emergency notes, and copies of specialist assessments of deliberate self poisoning. We checked all inpatient data retrospectively against information on admission and discharge for deliberate self poisoning that we obtained from the patient administration system in each hospital. We collected demographic details of patients, along with details of substance dependence, previous overdoses, and contact with psychiatric services. We also recorded information on the management of the current episode of self poisoning.

During the study period 458 patients accounted for 477 hospital attendances for deliberate self poisoning; 223 (49%) of these were women. The mean age of the patients was 30.9 years (SD 11.8 years); 65 (14%) were dependent on alcohol or drugs, 177 (39%) had taken a previous overdose, and 119 (26%) were in contact with psychiatric services. These percentages and the substances ingested were similar across study centres. By contrast, there were striking variations in the management of episodes between study centres, with a fourfold difference in discharge rates from accident and emergency departments, and almost a twofold difference in the proportion of subjects receiving a specialist psychosocial assessment (table). In 220 out of 477 hospital attendances (46%) the patient had no psychosocial assessment at any time during their hospital contact.

Comment

The average rate of patients with self poisoning presenting to hospital services in this study was 310 per 100 000 population per year, which suggests that deliberate self poisoning accounts for 170 000 hospital attendances in the United Kingdom annually. Yet services for this important problem remain in disarray. Striking variations in clinical practice were not accounted for by differences in patients' characteristics. We also discovered that, notwithstanding guidelines issued by the Department of Health,[4] almost half of the patients in this study did not receive a specialist psychosocial assessment.

Our findings may reflect a high risk approach to intervention or a lack of consensus on the psychiatric management of self poisoning.[5] We believe they probably reflect the low medical and psychiatric priority given to patients who have taken an overdose. A reduced number of beds means that medical staff are reluctant to admit patients who are judged to be at low physical risk and often seen as difficult and unrewarding. Meanwhile, psychiatric services are increasingly reserved for those with serious mental illness, a term which is not taken to include most cases of self poisoning. The current situation should not be allowed to continue because self poisoning represents a major social and clinical problem. At least, large scale intervention studies are required to inform practice and ensure that our management of deliberate self

Department of Liaison Psychiatry, Leeds General Infirmary, Leeds LS1 3EX
Navneet Kapur, *research registrar*
Allan House, *senior lecturer*
Department of Psychiatry and Behavioural Sciences, Manchester Royal Infirmary, Manchester M13 9WL
Francis Creed, *professor*
Elspeth Guthrie, *senior lecturer*
Department of Psychological Medicine, John Radcliffe Hospital, Oxford OX3 9DU
Eleanor Feldman, *consultant in liaison psychiatry*
Leicester General Hospital, Leicester LE5 4PW
Trevor Friedman, *consultant in liaison psychiatry*
Correspondence to: Dr N Kapur, Department of Psychiatry and Behavioural Sciences, Manchester Royal Infirmary, Manchester M13 9WL

BMJ 1998;316:831–2

Management of 477 episodes of deliberate self poisoning in each study centre. Values are numbers (percentages) of patients

Centre (No of episodes)	Discharged from accident and emergency department	Discharged from accident and emergency department without psychosocial assessment	Received psychosocial assessment	Admitted to psychiatric ward	Followed up*
Leeds (101)	18 (18)	15 (15)	65 (65)	6 (6)	46 (45), general practitioner 16 (16), deliberate self harm team 28 (28), psychiatric services 10 (10), alcohol services
Leicester (111)	61 (55)	23 (21)	76 (68)	18 (16)	50 (45), general practioner 12 (11), deliberate self harm team 44 (40), psychiatric services 2 (2), alcohol services
Manchester (100)	71 (71)	46 (46)	36 (36)	11 (11)	67 (67), general practitioner 0, deliberate self harm team 31 (31), psychiatric services 1 (1), alcohol services
Nottingham (165)	53 (32)	42 (25)	80 (48)	13 (8)	98 (59), general practitioner 14 (8), deliberate self harm team 43 (26), psychiatric services 6 (4), alcohol services
All centres (477)	203 (43)	126 (26)	257 (54)	48 (10)	261 (55), general practitioner 42 (9), deliberate self harm team 146 (31), psychiatric services 19 (4), alcohol services

*Nine episodes were followed up by a variety of agencies, mostly social services or non-statutory agencies.

poisoning in the future is less arbitrary than it has been for the past three decades.

We thank hospital staff at all four centres for their help with data collection, and Dr David Owens and an anonymous referee for their comments on earlier drafts of this paper.

Contributors: NK helped with the study design, collected and analysed all data, and contributed to writing the initial draft of the paper. AH helped to design the study and contributed to writing the initial draft of the paper. FC, EF, TF, and EG contributed to final aspects of study design, helped collect data, and commented on drafts of the paper. NK and AH are guarantors for the study.

Funding: This study was funded by the Medical Research Council, grant number G9533540, to AH, FC, Professor K Hawton, and Professor M Williams. At the time of this study NK was supported by the Medical Research Council and Leeds Community Mental Health Trust.

Conflict of interest: None.

1 Hawton K, Fagg J, Simkin S, Bale E, Bond A. *Attempted suicide in Oxford 1995*. Oxford: University Department of Psychiatry, 1996.
2 Hawton K, Fagg J. Suicide and other causes of death following attempted suicide. *Br J Psychiatry* 1988;152:359-66.
3 Hawton K, James R. General hospital services for attempted suicide patients: a survey in one region. *Health Trends* 1995;27(1):18-21.
4 Department of Health and Social Security. *The management of deliberate self-harm*. London: Department of Health and Social Security 1984. (HN(84)25.)
5 House A, Owens D, Storer D. Psycho-social intervention following attempted suicide: is there a case for better services? *Int Rev Psychiatry* 1992;4:15-22.

(Accepted 23 September 1997)

Survival rates from interval cancer in NHS breast screening programme

Stuart Collins, Ciaran B J Woodman, Anthony Threlfall, Pat Prior

Centre for Cancer Epidemiology, University of Manchester, Christie Hospital NHS Trust, Manchester M20 4QL
Stuart Collins, *information scientist*
Ciaran BJ Woodman, *professor of public health and cancer epidemiology*
Anthony Threlfall, *research officer*
Pat Prior, *research fellow*

Correspondence to: Mr Collins

BMJ 1998;316:832–3

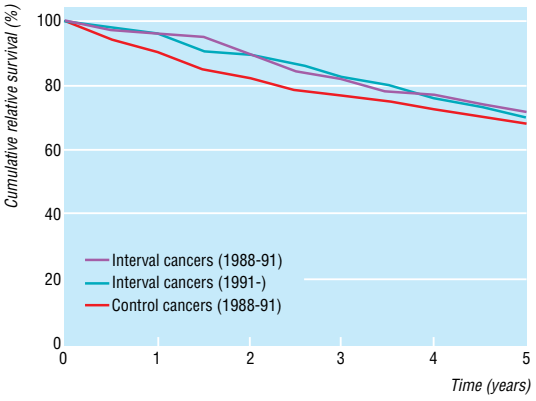
The NHS breast screening programme invites women aged 50-64 for screening every 3 years. In this programme the term interval cancer is applied to a breast cancer occurring within 3 years of a screening test with negative results. Substantially higher than anticipated rates of interval cancers have already been reported from the NHS breast screening programme,[1 2] and there is conflicting evidence on whether the survival rates of women with interval cancers are different from those of women with breast cancer occurring in an unscreened population.[3 4] Were interval cancers to have a worse prognosis than cancers in an unscreened population, the reduction in mortality from breast cancer in the screened population might be substantially less than predicted.

To interpret survival estimates for women with interval cancers requires identification of a suitable group of unscreened women for comparison. In the context of a national screening programme this is difficult. Women who do not respond to an invitation for screening, for example, have been shown to have a worse outcome than unscreened women and are therefore unsuitable.[4] The use of historical controls may also be inappropriate because of recent advances in managing breast cancer. Fortuitously, the phased introduction of the NHS screening programme in the north west has resulted in a group of women with breast cancer who lived in areas where screening had yet to be introduced whose survival can be compared with that of women diagnosed with interval cancers during the same calendar period. We report for the first time survival rates for interval cancers diagnosed during 1988-91 in the NHS breast screening programme.

Subjects, methods, and results

The NHS breast screening programme in the north west started in 1988 and by 1991 was under way in 14 district health authorities. Women resident in the five remaining districts in the region were not invited for screening before 1991 and form the control population. We identified all invasive interval cancers diagnosed between 1988 and 1991, using published methods.[1] We identified breast cancers presenting during this period in women aged 50-67 years in the control population from records held by the regional cancer registry. We determined date of death from data routinely notified to the registry. We calculated estimates of relative survival over five years and made comparisons using an appropriate proportional hazards regression that controlled for age.[5]

Seventy three interval cancers and 565 cancers from the control population were diagnosed during the study period. No significant difference could be shown between the relative survival rates of women from the control population with breast cancer and those of women presenting with interval cancers during the same period (hazards ratio 0.81 (95% confidence interval 0.50 to 1.31), $\chi = 0.67$, df = 1, P = 0.41). The robustness of this result is supported by the analysis of a further 441 interval cancers diagnosed after 1991, which showed a survival curve similar to that based on the 73 cancers (figure).



	No of women at risk in interval (years)				
	0-1	1-2	2-3	3-4	4-5
Interval cancers (1988-91)	73	70	65	59	55
Interval cancers (1991-)	439	389	275	141	42
Control cancers (1988-91)	565	509	460	427	398

Relative survival rates for women with interval cancers and women with breast cancer in control group

Comment

Although our results suggest no difference between the survival rates of women with interval cancers and those of women from the control population, variations in the quality of care provided for women from the two distinct areas could have invalidated this comparison. However, an analysis of survival rates for breast cancer, undertaken for the period immediately before the introduction of the screening programme in the north west, showed no significant differences when women were grouped according to their district of residence. It is reassuring that breast screening has not been detrimental to the survival of those women who presented with an interval cancer in the NHS screening programme. However, minimising the occurrence of interval cancers must remain a high priority if substantial reductions in mortality are to be achieved.

Contributors: SC participated in the discussion of core ideas and in identifying the study population, carried out data verification, undertook the survival analyses and interpretation of the results, and contributed to writing the paper. CBJW initiated and coordinated the research, participated in the discussion of core ideas, helped in interpreting the results, and contributed to the writing of the paper. AT initiated the research, participated in the discussion of core ideas, identified and collated data on the study population, and contributed to the writing of the paper. PP participated in the discussion of core ideas, helped in interpreting the results and contributed to the writing of the paper. All authors are guarantors for the study.

Funding: None.

Conflict of interest: None.

1 Woodman CBJ, Threlfall AG, Boggis CRM, Prior P. Is the three year breast screening interval too long? Occurrence of interval cancers in NHS breast screening programme's north western region. *BMJ* 1995;310:224-6.
2 Day N, McCann J, Camilleri-Ferrante C, Britton P, Hurst G, Cush S, et al. Monitoring interval cancers in breast screening programmes: the East Anglian experience. *J Med Screening* 1995;2:180-5.
3 Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F, et al. Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. *BMJ* 1988;297:943-8.
4 Tabar L, Faggerberg G, Duffy SW, Day NE, Gad A, Grontoft O. Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* 1992;30:187-210.
5 Hakulinen T, Abeywickrama KH. A computer program package for relative survival analysis. *Computer Programs in Biomedicine* 1985;19:197-207.

(Accepted 22 October 1997)

One hundred years ago

Wanted: brains

Dr Burt G Wilder, Professor of Physiology on the Cornell Staff of Instruction at Ithaca, has recently issued a circular asking prominent men in the United States to bequeath their brains to the university. He says that while it is easy to procure the brains of criminals and of insane or ignorant persons, it has hitherto been extremely difficult to obtain those of persons in whom the cerebral development is beyond the average. He adds that it is highly desirable for the advancement of science that a considerable number of brains of this character should be secured. This request, which has been circulated principally among the students and graduates of Cornell, is accompanied by a blank form of bequest, which, however, contains a clause by which the legacy becomes void if serious objection is made by the relatives of the deceased. (*BMJ* 1898;ii:1359)