Comparison of physiotherapy, manipulation, and corticosteroid injection for treating shoulder complaints in general practice: randomised, single blind study

Jan C Winters, Jan S Sobel, Klaas H Groenier, Hans J Arendzen, Betty Meyboom-de Jong

Department of General Practice, University of Groningen, Ant Deusinglaan 4, 9713 AW Groningen, Netherlands
Jan C Winters, *general practitioner*
Jan S Sobel, *general practitioner*
Klaas H Groenier, *statistician*
Betty Meyboom-de Jong, *professor of general practice*

Department of Rehabilitation Medicine, University Hospital Groningen, 9700 RB, Groningen, Netherlands
Hans J Arendzen, *medical director of clinic for rehabilitation medicine "Beatrixoord"*

Correspondence to: Dr J C Winters, Nieuwe Schoolweg 2A, 9756 BB Glimmen, Netherlands.

BMJ 1997;314:1320–5

Abstract

Objective: To compare the efficacy of physiotherapy, manipulation, and corticosteroid injection for treating patients with shoulder complaints in general practice.
Design: Randomised, single blind study.
Setting: Seven general practices in the Netherlands.
Subjects: 198 patients with shoulder complaints, of whom 172 were divided, on the basis of physical examination, into two diagnostic groups: a shoulder girdle group (n = 58) and a synovial group (n = 114).
Interventions: Patients in the shoulder girdle group were randomised to manipulation or physiotherapy, and patients in the synovial group were randomised to corticosteroid injection, manipulation, or physiotherapy.
Main outcome measures: Duration of shoulder complaints analysed by survival analysis.
Results: In the shoulder girdle group duration of complaints was significantly shorter after manipulation compared with physiotherapy ($P < 0.001$). Also the number of patients reporting treatment failure was less with manipulation. In the synovial group duration of complaints was shortest after corticosteroid injection compared with manipulation and physiotherapy ($P < 0.001$). Drop out due to treatment failure was low in the injection group (17%) and high in the manipulation group (59%) and physiotherapy group (51%).
Conclusions: For treating shoulder girdle disorders, manipulation seems to be the preferred treatment. For the synovial disorders, corticosteroid injection seems the best treatment.

Introduction

In the Netherlands most patients with shoulder complaints are diagnosed and treated by their general practitioner.[1] However, little research has been done to evaluate the effect of the different treatments given. The only trials of the efficacy of manipulation that we could find concerned treating the cervical spine. Among systematic reviews on the effect of non-steroidal anti-inflammatory drugs, corticosteroid injection, and physiotherapy on shoulder complaints, we found only three studies of treatment in general practice.[2-4] Vecchio *et al* compared a injection of corticosteroid into the subacromial bursa with injection of local anaesthetic and found no difference in efficacy.[5] Similarly, Jonquière found no difference in the effect of "classic" physiotherapy and that of "Cyriax" physiotherapy.[6] However, Lacey *et al* did find that treatment with a non-steroidal anti-inflammatory drug was significantly better than placebo.[7]

In view of the large number of patients with shoulder complaints who are treated in general practice and the lack of studies evaluating different treatments, we set up a trial to find the most effective treatment of shoulder complaints in general practice. Thus, we compared the effects of physiotherapy, corticosteroid injection, and manipulation.

Patients and methods

Patients

Between September 1994 and September 1995, all patients who consulted seven general practices in the Netherlands with shoulder complaints were included in our study unless one of the exclusion criteria applied. The study was approved by the ethics committee of the Medical Department of the University of Groningen. Before participating in the study, the patients had to give written informed consent.

Shoulder complaints were defined as pain localised in the region of the deltoid muscle, acromioclavicular joint, superior part of the trapezoid muscle, and scapula. Radiation of the pain in the arm could be present, and, besides the pain, the range of movement of the upper arm or shoulder girdle could be limited.

Exclusion criteria were treatment for shoulder complaints in the six months before consultation; bilateral shoulder complaints; presence of specific rheumatic disorders (polymyalgia rheumatica, rheumatoid arthritis, systemic lupus erythematosus, and fibromyalgia); shoulder complaints because of acute severe trauma such as fracture, dislocation, and cuff rupture (patients with a history of minor trauma were not excluded); presence of herniated cervical disc; presence of dementia or other psychiatric disorders; and refusal.

Allocation to treatment

On entry to the study, the patients' level of pain was established and they underwent a physical examin-

ation. On the basis of these the patients were allocated to three diagnostic groups: a synovial group, a shoulder girdle group, and a group with combinations of synovial and shoulder girdle disorders. For the first week, all patients were prescribed diclofenac sodium 50 mg thrice daily. At the end of the week, the patients' level of pain was measured again and the physical examinations were repeated by the general practitioners.

On the basis of this second diagnosis, patients were divided into two diagnostic groups: a shoulder girdle group and a synovial group (which also included the combination group because a previous study had shown that the course of complaints of the combination group and the synovial group was the same[8]). Randomisation to treatment took place separately in these two groups: patients in the synovial group were randomised to corticosteroid injection, manipulation, or physiotherapy, while those in the shoulder girdle group were randomised to manipulation or physiotherapy (injections could not be given in this group).

Assessment

Pain measurement—The severity of the shoulder complaints was assessed with the shoulder pain score, which is a six item questionnaire together with a 101 point numerical pain scale (for the total experienced pain).[9] The six questions—pain at rest, pain during motion, pain during the night, sleeping problems because of pain, inability to lie on the affected side, and presence of radiated pain—were scored on a four point scale of severity. The score on the 101 point numerical pain scale was also converted to a four point scale in order to calculate the sum score of the shoulder pain score. The range was from 7 points (no pain) to 28 points (severe pain).[9]

Physical examination consisted of measuring the active and passive range of movement of the glenohumeral joint, cervical spine, and upper thoracic spine and palpating the muscle tendons on the head of the humerus, the acromioclavicular joint, and the upper ribs. The examinations on inclusion into the study and before randomisation to treatment were performed by the seven participating general practitioners. Follow up examinations were done by a physiotherapist. In order to limit variation between doctors, the researchers had several sessions practising the physical examination and diagnostic interpretation.

Diagnostic groups

The three diagnostic groups have been described in detail elsewhere.[8]

The synovial group consisted of patients with pain or limited movement in one or several directions of the glenohumeral joint. These complaints originated from disorders of the subacromial structures, the acromioclavicular joint, the glenohumeral joint, or combinations of these (the synovial structures).

The shoulder girdle group consisted of patients with pain and sometimes slightly limited range of active movement of the glenohumeral joint. These problems were not related to the synovial structures but, instead, probably originated from functional disorders of the cervical spine, upper thoracic spine, or the upper ribs (the shoulder girdle).

The combination group consisted of patients with pain and sometimes slightly limited range of active or passive movement of the glenohumeral joint together with pain or limited range of movement of the cervical spine, upper thoracic spine, or upper ribs. Both the synovial structures and the structures of the cervical spine, upper thoracic spine, or upper ribs could have caused these complaints.[8]

Treatment

Corticosteroid injection consisted of an injection of 1 ml of 40 mg/ml triamcinolone acetonide in combination with 9 ml of 10 mg/ml lignocaine. One to three injections were given by the participating doctors immediately after randomisation, one week later, and, if needed, after a further two weeks. In each treatment session two out of the three synovial structures (glenohumeral joint capsule, subacromial space, and acromioclavicular joint) were injected. We chose this multiple injection scheme because most of the patients in the synovial group had combinations of disorders of the synovial structures.[8] Using a multiple injection scheme modified from that of Steinbroker *et al*,[10] Roy *et al* had successfully treated frozen shoulder.[11] Our injection techniques were standardised: the intra-articular injection was given from the posterior side, the subacromial injection from the lateral side, and the acromioclavicular injection perpendicularly from the upper side of the joint.

Physiotherapy was given twice a week by local physiotherapists. They were instructed to use "classic" physiotherapy—such as exercise therapy, massage, and physical applications. No mobilisation techniques or manipulative techniques were allowed. This definition of physiotherapy was satisfactorily used by Koes *et al* in their study of treating low back pain.[12]

Manipulation consisted of mobilisation and manipulation of the cervical spine, upper thoracic spine, upper ribs (on the segmental level), acromioclavicular joint, and the glenohumeral joint once a week with a maximum of six treatment sessions. The manipulation was done by either the participating general practitioners or physiotherapists (graduates from the Eindhoven course for manipulative therapy). They were instructed in which techniques to use.

Follow up

After treatment had started, the patients weekly filled in the pain questionnaire. They were also asked to indicate if they felt "cured" or if the treatment failed. Feeling cured was defined as disappearance of shoulder complaints or a decrease of shoulder complaints to such an extent that they were no longer inconvenient, did not need treatment, or no longer interfered with normal working. In our previous study of the pain questionnaire we found that patients did not need to be totally free of pain to feel "cured."[9] Treatment failed when a patient experienced no improvement or the condition deteriorated.

Follow up examinations were done by a physiotherapist at two, six, and 11 weeks after randomisation. If a patient felt cured or the treatment had failed a final examination was done as soon as possible. At the end of the study the physiotherapist contacted these patients to inquire about present complaints. If patients did have complaints, their level of

Table 1 Characteristics of cohort of patients in Netherlands who consulted their general practitioners for shoulder complaints. Values are numbers (percentages) unless stated otherwise

Characteristics	All patients enrolled (n=198)	Patients dropped out after NSAID treatment (n=25)
Mean (SD) age (years)	49.3 (12.9)	48.9 (14.4)
Women	111 (56)	13 (52)
Right handed	176 (89)	22 (88)
Married or living with partner	162 (82)	20 (80)
History of previous shoulder complaints	86 (43)	10 (40)
History of minor trauma	26 (13)	2 (8)
Employment:		
Full time	57 (29)	6 (25)
Part time	37 (19)	6 (25)
None	104 (52)	12 (50)
Light manual work with hands above shoulder level	51/94 (54)	3/12 (25)
Sickness absenteeism	15/94 (16)	4/12 (33%)
Problems in left shoulder:	86 (43)	7 (28)
Dominant side	13 (15)	5 (71)
Non-dominant side	73 (85)	2 (29)
Problems in right shoulder:	112 (57)	18 (72)
Dominant side	103 (92)	17 (94)
Non-dominant side	9 (8)	1 (6)
Duration of complaints before first consultation (weeks):		
≤1	45 (23)	11 (44)
2-4	51 (26)	5 (20)
5-25	64 (32)	6 (24)
≥26	38 (19)	3 (12)
Mean (SD) pain score on enrollment	18.7 (4.3)	17.8 (5.2)
Diagnostic group:		
Synovial	104 (53)	14 (56)
Shoulder girdle	46 (23)	4 (16)
Combination	48 (24)	7 (28)

NSAID=Non-steroidal anti-inflammatory drug.

pain was established and they could indicate whether they felt cured.

Outcome parameters and statistical analysis

Before the study began, power calculation showed that, with $\alpha = 0.05$ and a power of 80%, a difference of 0.8 standard deviation could be detected in treatment groups of 25 patients. Our aim was to achieve treatment groups of at least 30 patients.

To evaluate the effect of treatment, we analysed the duration of patients' complaints, treatment failures, and any complaints at the end of the study on an intention to treat basis. We analysed the duration of the complaints with a survival analysis (log rank test), also known as event history analysis. In this study the event we studied was patients' feeling "cured," and we corrected the calculations for patients who dropped out because of treatment failure. We evaluated the differences between group averages with analysis of variance or Student's t test and analysed the difference between group numbers with the χ^2 test.

Assignment

The university's Department of Family Practice was in charge of the randomisation to treatment. For each diagnostic category, we had made a series of closed unnumbered envelopes which contained instructions of the treatment to be given. The participating general practitioners had to call a secretary and state the diagnostic category of each patient. The secretary in turn would draw an envelope to assign treatment.

Masking (blinding)

The follow up examinations after randomisation were done by a physiotherapist who was not informed about the patients' diagnosis and treatment.

Results

A total of 198 patients enrolled in the study, and table 1 summarises their characteristics. A substantial proportion of the patients had previously experienced shoulder complaints, and almost 20% had had shoulder complaints for six months or more before consultation. About half of the patients had a synovial syndrome, a quarter had a shoulder girdle syndrome, and a quarter had a combination syndrome.

Twenty six patients dropped out of the study before randomisation. One patient dropped out because of family circumstances. The other 25 considered themselves to be cured after the week's treatment with non-steroidal anti-inflammatory drug: their mean pain scores had dropped from 17.8 to 8.9. These patients were generally similar to the total cohort enrolled, though they had a shorter history of shoulder complaints (table 1).

Figure 1 shows how the remaining 172 patients were randomised to treatment. The increased number of patients in the shoulder girdle group (from 46 patients to 58) was because of a diagnostic shift from the synovial or combination group towards the shoulder girdle group as a result of the treatment with non-steroidal anti-inflammatory drug.

Table 2 shows the characteristics of these groups. The shoulder girdle group was younger than the synovial group.

Table 2 Characteristics of 172 patients with shoulder complaints at randomisation to treatment. Values are numbers of patients unless stated otherwise

	Shoulder girdle group (n=58)		Synovial group (n=114)		
Characteristic	Manipulation (n=29)	Physiotherapy (n=29)	Corticosteroid injection (n=47)	Manipulation (n=32)	Physiotherapy (n=35)
Mean (SD) age (years)	43.9 (12.6)	46.4 (11.2)	53.5 (12.5)	46.7 (12.1)	53.1 (12.6)
Women	15	18	32	17	14
History of minor trauma	2	4	8	2	8
Median duration of complaints before first consultation (weeks)*	3	4	8	9	4
History of previous shoulder complaints	14	10	27	8	16
Mean (SD) pain score at randomisation	14.9 (4.2)	14.5 (3.5)	16.3 (4.8)	15.7 (4.2)	16.3 (3.8)
Paid employment	14	16	32	13	15

*Duration of complaints measured up to maximum of 25 weeks, after which it was simply recorded as 26 weeks.

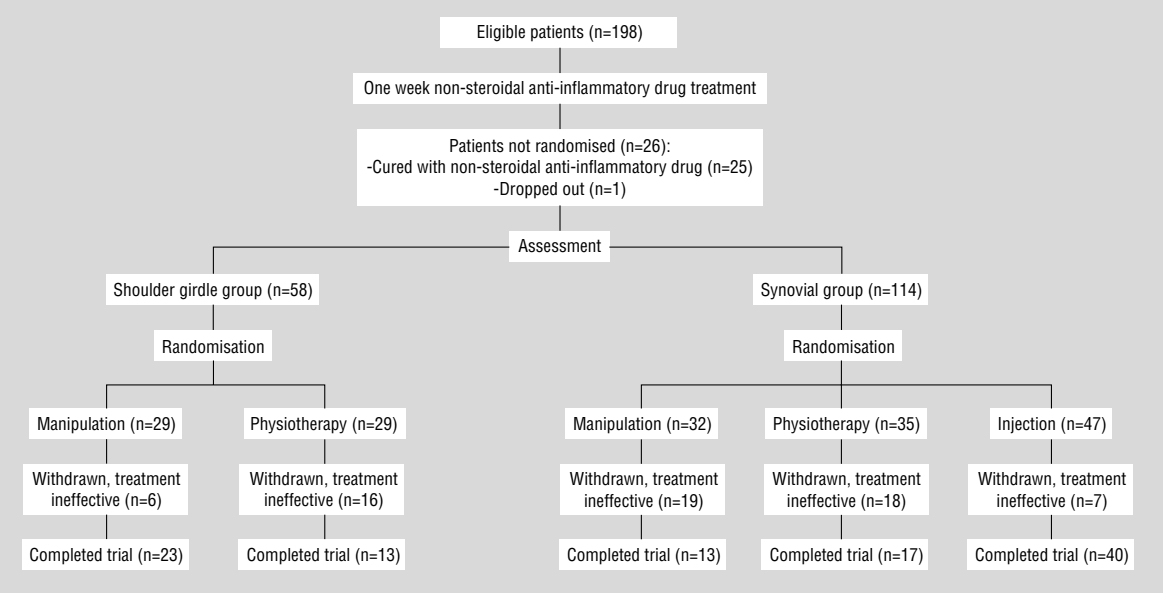

Fig 1 Flow chart describing progress of patients through trial

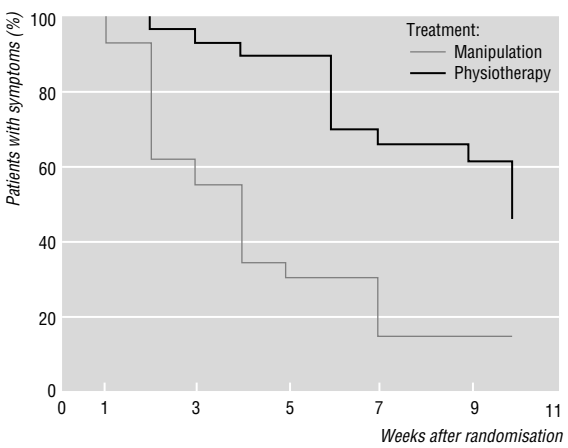

Fig 2 Duration of shoulder complaints in shoulder girdle group after randomisation to treatment

Shoulder girdle group

Figure 2 shows the survival analysis of the shoulder girdle group. Manipulation was superior to physiotherapy ($P<0.001$): at five weeks after randomisation almost 70% of the patients in the manipulation group considered themselves to be cured compared with 10% of the physiotherapy group. Drop out because of treatment failure was significantly higher in the physiotherapy group (45% (13/29) of patients) than in the manipulation group (20% (6/29) of patients).

Table 3 shows the two treatment groups' pain scores at randomisation and the final pain scores (on being "cured" or at 11 weeks after randomisation). Both treatments significantly reduced the patients' pain scores. When we differentiated between patients who were "cured" and those who were not, we found that the reductions in the pain scores in both treatment groups were due to the "cured" patients. Of the patients who were "cured" before week 11 after randomisation, 15% (2/13) of patients in the physiotherapy group and 9% (2/22) of patients in the manipulation reported a recurrence of complaints by week 11 after randomisation.

Synovial group

Figure 3 shows the survival analysis of the three treatment groups in the synovial group. The corticosteroid injection group (average number of injections was 1.8) improved rapidly, while the physiotherapy group improved slowly and the manipulation group did only slightly better ($P<0.001$): at five weeks after randomisation, 75% of patients in the injection group were "cured" compared with 20% in the physiotherapy group and 40% in the manipulation group. Drop out because of treatment failure was much lower in the injection group (17% (7/47)) than in the physiotherapy group (51% (18/35)) and manipulation group (59% (19/32)).

Table 3 shows that all three treatments significantly reduced the patients' pain scores. Again, the patients

Table 3 Mean (SD) pain scores of patients with shoulder complaints at randomisation to treatment and at end of treatment (when patient left study or 11 weeks after randomisation)

	Shoulder girdle group (n=58)		Synovial group (n=114)		
Pain score	Manipulation (n=29)	Physiotherapy (n=29)	Corticosteroid injection (n=47)	Manipulation (n=32)	Physiotherapy (n=35)
At randomisation	14.8 (4.2)	14.4 (3.5)	16.3 (4.8)	15.7 (4.2)	16.3 (3.3)
At end of treatment:	9.9 (3.5)†	12.0 (4.4)†	9.2 (3.7)†	12.6 (5.1)†	11.5 (4.4)†
Patients who were "cured"*	8.7 (2.2)†	9.6 (2.7)†	8.3 (1.7)†	8.9 (1.4)†	8.2 (1.5)†
Patients who were not cured	14.3 (4.1)	14.9 (4.4)	13.5 (6.9)	15.0 (5.1)	14.6 (3.8)

*See text for details.
†Significant difference from pain score at randomisation ($P\leq0.001$).

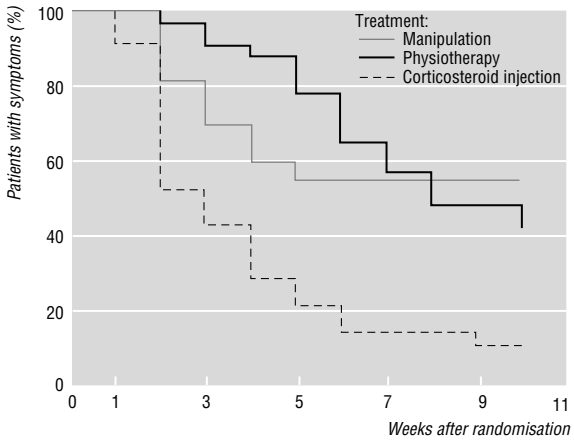

Fig 3 Duration of shoulder complaints in synovial group after randomisation to treatment

who were "cured" accounted for this reduction. Of the patients who were "cured" before 11 weeks after randomisation, a recurrence of complaints by week 11 was reported by 18% (7/39) of patients in the injection group 13% (2/15) in the physiotherapy group, and 8% (1/13) in the manipulation group.

Discussion

Design of study

Our study design was based on the results of our earlier descriptive study.[8] During that study it became evident that other diagnostic classifications, such as those by Cyriax[13] and the National Guidelines for Shoulder Complaints of the Dutch College of General Practitioners,[14] were not suitable for diagnosing shoulder complaints in general practice. Shoulder complaints seem to be often due to problems in various structures in and around the glenohumeral joint or the structures of the shoulder girdle.

Patients were prescribed a non-steroidal anti-inflammatory drug in the first week after enrollment in order to reduce moderate to severe pain to light to moderate pain. This allowed us to treat patients with physiotherapy and manipulation without having to give additional treatment for the pain. In our study 13% of the patients were "cured" after the non-steroidal anti-inflammatory drug treatment.

Despite the randomisation procedure, in the synovial group the patients allocated to manipulation were significantly younger than the patients allocated to the two other treatments, and the percentage of men in the physiotherapy group was significantly higher than in the other groups. In a separate regression analysis we concluded that sex did not have a significant influence on the duration of complaints but the age of patients did. Thus, the lower age of the patients given manipulation group could have influenced the better results that they showed in the first 6 weeks after randomisation compared with the physiotherapy group. However, the results of manipulation in the group were modest, especially when compared with the results of manipulation in the shoulder girdle group, which had the same average age.

Key messages

- Many patients with shoulder complaints are treated in general practice, but there has been little evaluation of different treatments
- In this single blind randomised trial we investigated the effect of corticosteroid injection, manipulation, and physiotherapy on the duration of shoulder complaints among patients treated in general practice
- Patients were divided into two diagnostic groups: those with complaints originating from the synovial structures and those whose complaints originated from the shoulder girdle
- The first group was treated by injection, manipulation, or physiotherapy: those receiving injection showed quickest recovery and only 17% of patients dropped out because of treatment failure, compared with 51% in physiotherapy group and 59% in manipulation group
- Patients with a shoulder girdle disorder were treated by manipulation or physiotherapy: duration of complaints was significantly shorter after manipulation, and there was only 20% drop out in this group compared with 45% in physiotherapy group

Implications of results

To our knowledge, no other published study has described the positive effects of manipulation in treating shoulder complaints. The results of our study suggest that manipulation is to be preferred to physiotherapy for treating shoulder complaints originating from the shoulder girdle in general practice.

Of 22 comparative studies that investigated corticosteroid injection for treating shoulder complaints, only five describe success with injection.[3] We consider our positive results with corticosteroid injection were helped by our setting in general practice (no patient selection by referral) and adequate selection of patients by diagnostic groups. We found injection to be the most effective treatment for shoulder complaints originating from the synovial structures in general practice (after a week's treatment with a non-steroidal anti-inflammatory drug). A slightly higher percentage of the "cured" patients in the injection group reported recurrence of complaints at the end of the study. However, in this group 80% of the patients were "cured" by the fifth week after randomisation so these patients had the longest period for symptoms to recur.

We thank our colleague general practitioners Luit-Jan Lukkes, Mello Maaskant, Henk Spelde, Jan de Weerd, and Jan Woudhuizen for their willingness to participate in this study.

Funding: This study was conducted with a grant of the Ministry of Welfare, Health and Culture.

Conflict of interest: None.

1 Lamberts H, Wood M, Hofmans-Okkes JM. *The international classification of primary care in the European community*. Oxford: Oxford University Press, 1993.

2 Winters JC, Sobel JS, Arendzen JH, Meyboom-de Jong B. NSAID's bij schouderaandoeningen [NSAID's and shoulder complaints]. *Huisarts en Wetenschap* 1995;38:339-41, 347.

3 Winters JC, Sobel JS, Arendzen JH, Meyboom-de Jong B. Het effect van corticosteroïdinjecties bij schouder-aandoeningen [The effect of corticosteroid injections on shoulder complaints]. *Huisarts en Wetenschap* 1995; 38:164-8.
4 Van der Heijden GJMG, Bouter LM, Beckerman H, de Bie RA, Oostendorp RAB. De effectiviteit van fysiotherapy bij schouderklachten, een geblindeerd literatuuronderzoek [The effect of physiotherapy on shoulder complaints, a blinded literature research]. *Nederlands Tijdschrift voor Fysiotherapie* 1992;102:38-46.
5 Vecchio PC, Hazleman BL, King RH. A double-blind trial comparing subacromial methylprednisolone and lignocaine in acute rotator cuff tendinitis. *Br J Rheumatol* 1993;32:743-5.
6 Jonquière M. De behandeling van schouder-en elleboogaandoeningen volgens de richtlijnen van Cyriax [The treatment of shoulder and elbow complaints according to the guidelines of Cyriax] [dissertation]. Rotterdam: Erasmus University, 1986: 121-35.
7 Lacey PH, Dodd GD, Shannon DJ. A double blind, placebo controlled study of piroxicam in the management of acute musculoskeletal disorders. *Eur J Rheumatology Inflammation* 1984;7:95-104.
8 Sobel JS, Winters JC, Arendzen JH, Groenier KH, Meyboom-de Jong B. Schouderklachten in de huisartspraktijk [Shoulder complaints in general practice]. *Huisarts en Wetenschap* 1995;38:342-7.
9 Winters JC, Sobel JS, Groenier KH, Arendzen JH, Meyboom-de Jong B. A shoulder pain score; a comprehensive questionnaire to assess pain in patients with shoulder complaints. *Scan J Rehabil Med* 1996;28:163-7.
10 Steinbroker O, Thomas G, Argyros MD. Frozen shoulder: treatment by local injections of depot corticosteroids. *Arch Phys Med Rehabil* 1974;55: 209-12.
11 Roy S, Oldham R. Management of the frozen shoulder. *Lancet* 1976;i: 1322-4.
12 Koes BW, Bouter LM, van Mameren H, Essers AHM, Verstegen GMR, Hofhuizen DM, *et al*. The effectiveness of manual therapy, physiotherapy and treatment by the general practitioner for non-specific back and neck complaints: a randomised clinical trial. *Spine* 1992;17:28-35.
13 Cyriax J. *Textbook of orthopaedic medicine*. 11th ed, vol 1. *Diagnosis of soft tissue lesions*. London: Bailliere Tindall, 1984: 127-42, 143-58.
14 Bakker JF, de Jongh AC, Jonquière M, Mens J, Oosterhuis WW, Poppelaars A, *et al*. NHG Standaard Schouderklachten [National Guidelines for Shoulder Complaints of the Dutch College of General Practitioners]. *Huisarts en Wetenschap* 1990;33:196-202.

(Accepted 24 January 1997)

The social origins of infantile colic: questionnaire study covering 76 747 infants

N S Crowcroft, D P Strachan

Abstract

Objective: To describe risk factors for infantile colic.
Design: Questionnaire administered by health visitors.
Setting: Sheffield.
Subjects: Mothers of 76 747 infants born between 1 August 1975 and 31 May 1988, interviewed when the infant was 1 month old.
Main outcome measures: Reporting of infantile colic and its duration; weight of infant, feeding, state of the home, socioeconomic characteristics of the parents, parents' age, and mother's parity.
Results: The odds of reporting infantile colic were increased with breast feeding (odds ratio of breast *v* bottle feeding 1.35 (95% confidence interval 1.28 to 1.43)), increasing parental age, lower parity, increasing parental age at leaving full time education, and more affluent homes and districts of residence. In a logistic regression analysis, mother's age and parity and socioeconomic factors remained the most important risk factors for the reporting of infantile colic (each $P < 0.005$), and the effect of breast feeding was attenuated (odds ratio of breast *v* bottle feeding 1.09 (1.02 to1.15)).
Conclusion: At a population level, dietary factors contribute little to mothers' reporting of infantile colic, and dietary change should not be the primary intervention.

Introduction

Infantile colic is a syndrome characterised by paroxysmal, excessive, and inconsolable crying without identifiable cause in a healthy infant. It is also called persistent crying in infancy and three month colic because it usually disappears by three months of age. Infantile colic is common.[1] Estimates of cumulative incidence have varied, depending on the case definition and the period of follow up, from 10% to 40%.[2 3] It is usually self limiting, without long term adverse consequences, but caring for an infant with colic can be distressing and frustrating for parents. At the extreme its effects on the parent-infant relationship may be sufficient to disrupt the infant's development,[4-6] and it may increase the risk of child abuse.[7] Research has not established the aetiology or best management of infantile colic, and infant crying was identified as a priority area in the NHS Research and Development Programme.[8]

The Sheffield child development study was designed to identify infants at high risk of sudden infant death but included questions on parental reporting of colic.[9 10] We aimed to identify risk factors for parental reporting of infantile colic during the first month of life. Clarifying the aetiology may indicate possible approaches for prevention and for clinical management.

Subjects and methods

In the Sheffield child development study, all parents of infants born in Sheffield between 1975 and 1995 were interviewed by a health visitor at the routine visit when the infant was about one month of age. Our study used information collected from 1975 to 1988 because data were computerised only for this period.

A detailed questionnaire was completed on a range of subjects which included characteristics of the family, feeding practices, and the behaviour of the infant. Ascertainment of colic relied on parental reporting, as in other studies.[3 11] Health visitors asked mothers whether or not the baby was currently colicky as well as the number of days the baby had been colicky since birth. The questions on colic formed part of a short series of questions within the questionnaire which were all in the same format and which concerned various minor symptoms. Colic was not formally defined for

Department of Public Health Sciences, St George's Hospital Medical School, London SW17 0RE
N S Crowcroft, *senior registrar in public health medicine*
D P Strachan, *reader in epidemiology*
Correspondence to: Dr N S Crowcroft, Epidemiology Department, Scientific Institute of Public Health—Louis Pasteur, 14 Juliette Wytsmanstreet, Brussels B-1050, Belgium.

BMJ 1997;314:1325–8

Table 1 Birth weight, weight gain, and weight at one month in colicky and non-colicky babies. Values are means (95% confidence intervals)

Weight (grams)	Ever colicky	Never colicky	Difference (unadjusted)
Birth weight	3359.7 (2419 to 4301) n=12 271	3329.2 (2346 to 4312) n=54 849	30.5 (21 to 40) P<0.0001
Weight at one month	4091.1 (3035 to 5148) n=11 290	4039.6 (2922 to 5157) n=49 961	51.6 (40 to 63) P<0.0001
Weight gain	731.7 (111 to 1352) n=11 277	709.6 (83 to 1337) n=49 843	22.1 (16 to 29) P<0.0001
Weight gain as percentage of birth weight	22.5 (1 to 44) n=11 277	22.0 (0.5 to 44) n=49 843	0.45 (0.2 to 0.7) P<0.0001

Table 2 Occurrence and duration of colic and method of feeding. Values are percentages (numbers)

	Breast only	Changed from breast to bottle	Breast and bottle	Bottle only
Occurrence	100 (22 350)	100 (12 121)	100 (7294)	100 (21 564)
Still colicky at one month	14.1 (3152)	12.9 (1561)	11.9 (867)	10.2 (2203)
Colicky during first month but not at one month	6.1 (1378)	6.9 (831)	6.1 (441)	5.4 (1174)
Never colicky	79.7 (17 820)	80.3 (9729)	82.1 (5986)	84.3 (18 187)
Duration*	100 (22 285)	100 (12 081)	100 (7271)	100 (21 516)
>13 Days	8.1 (1802)	7.2 (868)	6.8 (491)	5.6 (1209)
1-13 Days	12.0 (2663)	12.3 (1484)	10.9 (794)	9.9 (2120)
No colic	80.0 (17 820)	80.5 (9729)	82.3 (5986)	84.5 (18 187)

*Duration of colic was not recorded for some infants who were reported to be colicky.

the study, but the health visitor could offer clarification if the mother asked for it, as at any routine visit.

The age at interview ranged from two to more than 99 days. The analysis was restricted to babies aged between 24 and 37 days, representing 87.5% (67 172/76 747) of the births and 92% of those for which data on colic were available (67 172/72 995).

Our analysis explored hypotheses identified from the literature[1 2 7 11] relating to diet and social and economic factors and tested the null hypothesis that babies who are reported to be colicky are no different from babies who are not.

Univariate analyses were performed using SAS version 6.09.[12] Relations between categorical variables were examined using the χ^2 test or, for ordered categorical variables, the χ^2 test for trend. Stratified tests for trend were carried out in EpiInfo version 5. Continuous variables were compared by using the large sample normal test for the differences between means and Pearson's correlation coefficient.[13] The logistic regression was carried out in GLIM4.[14] Variables were included if there was a significant association with colic in the univariate analysis (at a significance level of 0.05) or if they were likely confounders. The dependent variable was a dichotomous measure of parental reporting of colic (current and past) versus no history of colic. Explanatory variables were tested in a forward stepwise regression analysis using a χ^2 test of heterogeneity or for trend where appropriate.

Results

In total, 12 277 infants were reported to have been colicky at some stage in the first month, a cumulative incidence in the first month of life of 18.3% (12 277/67 127), and 8251 infants were recorded as currently being colicky, a point prevalence of colic at one month of 12.3% (8251/67 127). Colic and sex of the infant were not related, with a prevalence of 18.4% in boys and 18.2% in girls. The prevalence of colic in babies described as "demanding" was 38.9% (3105/7983) compared with 15.5% (8942/57692) in those described as contented or never crying.

Colicky babies weighed more at birth, gained more weight, and weighed more at 1 month of age (table 1). After adjusting for maternal education, parity and maternal age the differences in birth weight (21.33 g, 95% confidence interval 11.47 g to 21.21 g) and weight gain as a percentage of birth weight (0.44%, 0.21% to 0.67%) were trivial although they remained significant ($P<0.0001$).

Information was recorded about the longest period that the baby was reported to spend continuously asleep or awake. Colicky babies had significantly shorter periods of "longest continuous time asleep" (mean difference 14.8 (95% confidence interval 13.0 to 16.7) minutes, $P<0.0001$) and longer periods of "longest continuous time awake" (29.1 (27.1 to 31.0) minutes, $P<0.0001$).

Babies who had been or were being breast fed were significantly more likely to be reported as colicky ($\chi^2=198.4$, df=6, $P<0.0001$) (table 2). The prevalence of past or current colic in exclusively breast fed babies was 20% compared with 16% in exclusively bottle fed babies, a relative risk of 1.62 (1.55 to 1.71). There was also a significant trend for breast fed babies to have a longer reported duration of colic than bottle fed babies (χ^2 test for trend comparing breast only and bottle only, $P<0.0001$) (table 2).

Of the infants who were bottle fed, colicky babies were more likely to have had a change in type of formula feed and to have had more changes in type of formula than non-colicky babies. Of the infants who had one or more changes in formula, colicky infants had a mean of 1.2 changes, compared with 1.1 for non-colicky babies ($P<0.0001$).

Table 3 shows that there was a non-linear relation between maternal age and past or present colic, the unadjusted prevalence of colic being highest among the offspring of mothers aged 30-34 years. The prevalence and odds of reported colic fell progressively with increasing parity.

Health visitors subjectively assessed some characteristics of the environment of the child on a five point scale, including the type of neighbourhood and the state of repair of the house, furnishings, and equipment. The prevalence of reported colic showed a trend of increasing colic with more affluent neighbourhood and better state of repair of the home ($P<0.0001$).

Educational achievement was examined as a proxy measure for current social class, as parental occupation was not recorded in the Sheffield child development study until 1983. Increased rates of reporting were observed in mothers who were older when they left full time education (table 3). From 1983 to 1988, the prevalence of reported colic was higher where mothers or fathers had "white collar" occupations (($\chi^2=166.3$, df=8, for mother's or father's occupation, $P<0.0001$) (table 4).

The logistic regression analysis was performed to control for confounding and to estimate the strength of the risk associated with significant explanatory variables. Two logistic regression models were created, one

using all the data from 1975 to 1988 (table 3) and the other including only data from 1983 to 1988, when parental employment was recorded (not shown). Excluding the missing data reduced the number of infants to 56 949/67 172 (85%) for the period 1975-88 and to 25 952/33 554 (77%) for the period 1983-8.

The odds ratios for current feeding method shown in table 3 use exclusive breast feeding as the reference category. Bottle feeding was associated with reduced odds of reporting colic compared with breast feeding, after controlling for potential confounders (odds ratio 0.92). In the 1983-8 data, the ratio of the odds of reporting colic by mothers who bottle fed to that of those who breast fed was 0.93 (0.85 to 1.02) maternal age, parity, education and occupation were controlled for. Changing from breast to bottle significantly increased the odds of reported colic in comparison to exclusive breast feeding in both models (odds ratio 1.13 (1.06 to 1.20) *v* 1.13 (1.02 to 1.25)).

After adjustment for mother's and father's age and educational achievement, mother's parity, and feeding method, a strong trend remained of greater odds of reported colic in houses judged to be in a better state of repair, but the independent effect was weak. Maternal education was a stronger independent correlate of reported colic than father's school leaving age (table 3). The relation between mother's age and infantile colic was strengthened after adjustment, with an increase in the odds of colic in older mothers (table 3). The trend of lower odds of colic with increasing parity was little affected by adjustment.

Table 3 Unadjusted and adjusted odds ratios for colic, 1975-88

Categories	No	Unadjusted odds ratios	Adjusted odds ratios*	P value of adjusted odds ratio
Feeding method				
Breast	22 350	1.00	1.00	<0.0005†
Bottle	21 564	0.74 (0.70 to 0.78)	0.92 (0.87 to 0.98)	
Both	7 294	0.86 (0.81 to 0.93)	0.95 (0.88 to 1.02)	
Changed	12 121	0.98 (0.92 to 1.03)	1.13 (1.06 to 1.20)	
Other	338	0.86 (0.63 to 1.16)	1.12 (0.82 to 1.51)	
Repair of housing				
Very poor	399	1.00	1.00	<0.0005‡
Poor	3 372	1.34 (0.89 to 2.03)	1.29 (0.85 to 1.97)	
Average	22 502	1.84 (1.24 to 2.73)	1.59 (1.05 to 2.40)	
Good	30 137	2.36 (1.59 to 3.49)	1.78 (1.18 to 2.69)	
Very good	10 342	2.70 (1.82 to 4.01)	1.79 (1.18 to 2.71)	
District or residence				
Very poor	462	1.00	1.00	<0.005‡
Poor	6 303	1.43 (1.02 to 2.02)	1.13 (0.79 to 1.61)	
Average	34 457	1.67 (1.19 to 2.34)	1.08 (0.76 to 1.54)	
Good	21 787	2.14 (1.53 to 3.00)	1.14 (0.79 to 1.63)	
Very good	3 735	2.71 (1.92 to 3.82)	1.33 (0.92 to 1.93	
Mother's education				
< 15 Years	2 169	1.00	1.00	<0.0005§
15 Years	18 763	1.39 (1.20 to 1.62)	1.26 (1.07 to 1.49)	
16-18 Years	33 156	1.59 (1.89 to 2.56)	1.36 (1.16 to 1.61)	
>18 Years	8 657	2.20 (1.89 to 2.56)	1.50 (1.26 to 1.79)	
None	983	0.53 (0.39 to 0.72)	1.75 (1.27 to 2.41)	
Student	409	1.77 (1.31 to 2.38)	1.20 (0.85 to 1.69)	
Father's education				
< 15 Years	2 220	1.00	1.00	<0.001§
15 Years	20 843	1.33 (1.16 to 1.53)	1.02 (0.87 to 1.20)	
16-18 Years	28 013	1.52 (1.32 to 1.74)	1.09 (0.93 to 2.75)	
>18 Years	9 417	2.06 (1.79 to 2.38)	1.19 (1.00 to 1.41)	
None	296	0.65 (0.42 to 1.03)	1.03 (0.64 to 1.67)	
Student	732	1.77 (1.41 to 2.24)	1.22 (0.93 to 1.59)	
Parity				
0	24 394	1.00	1.00	<0.0005‡
1	22 296	0.86 (0.81 to 0.90)	0.84 (0.79 to 0.88)	
2	11 272	0.81 (0.76 to 0.86)	0.79 (0.74 to 0.84)	
3	5 024	0.82 (0.75 to 0.89)	0.82 (0.75 to 0.90)	
4	2 161	0.77 (0.67 to 0.87)	0.81 (0.70 to 0.92)	
5	991	0.61 (0.49 to 0.75)	0.68 (0.54 to 0.84)	
>5	983	0.50 (0.39 to 0.63)	0.59 (0.46 to 0.75)	
Mother's age (years)				
<20	6 861	1.00	1.00	<0.0005‡
20-24	20 657	1.09 (0.99 to 1.19)	1.11 (1.01 to 1.21)	
25-29	23 215	1.37 (1.26 to 1.49)	1.33 (1.22 to 1.46)	
30-34	12 192	1.44 (1.32 to 1.58)	1.42 (1.28 to 1.57)	
35-39	3 454	1.25 (1.11 to 1.41)	1.35 (1.18 to 1.54)	
40-44	651	1.35 (1.07 to 1.70)	1.73 (1.36 to 2.21)	
>44	91	1.12 (0.52 to 2.41)	2.39 (1.09 to 5.24)	

*Adjusted for all the other variables shown.†Test for heterogeneity.‡Test for trend.§Test for trend excluding none and student.

Discussion

This large, population based survey confirms that infantile colic is commonly reported and associated with an increased chance of the mother finding the infant demanding and changing the pattern of feeding. However, feeding method did not emerge as an important determinant of the incidence or duration of colic after allowance for the stronger influences of maternal age, parity, and socioeconomic circumstances.

Methodological issues

As these data cover a virtually complete set of births between 1975 and 1988, the problems of selection bias that are seen in many other studies of infantile colic have been avoided.[15] Furthermore, the size of this study gives it the power to exclude with confidence all but the most subtle associations. Even small differences between groups that may not be clinically important and that may be the result of residual confounding are highly statistically significant.

Cases of colic that developed after the infant was older than 1 month were not ascertained, but in 90% of cases, colic starts in the first month of life.[16] In this study, most babies were seen close to 1 month of age, and babies who were seen before 24 days were excluded from the analysis, so most colicky babies should have been identified.[15 17] However, babies who present before 1 month of age may differ systematically from those who present at a later stage, in having more severe colic.[16]

Table 4 Parent's occupation (1983-8) and percentage (number) of infants with colic

Occupation	Reporting of colic by mother's occupation	Reporting of colic by father's occupation
White collar	23.6 (955/4043)	23.5 (1002/4271)
Semiclerical	20.7 (1658/8010)	20.2 (982/4869)
Blue collar	17.9 (1339/7462)	18.6 (1841/9910)
Unemployed	14.6 (397/2721)	15.0 (752/5019)
Student		
Higher education	24.2 (53/219)	23.4 (61/261)
A levels or below	15.2 (60/395)	18.4 (16/87)
Housewife/husband	13.7 (274/2001)	18.8 (3/16)
YOPS/YTS*	16.3 (44/270)	19.8 (21/106)
Not known	16.2 (192/1189)	16.6 (294/1771)
Total	18.9 (4972/26 310)	18.9 (4972/26 310)

*Youth training programmes.

Key messages

- Infantile colic is poorly defined but commonly reported and causes parents appreciable distress
- Often the first advice that parents receive is to change the infant's diet
- Slightly higher rates of reporting of infantile colic are found when infants are breast fed than bottle fed, so formulas based on cows' and hence allergy to cows' milk protein are unlikely to be important causes of infantile colic
- Social factors are most important, with older primigravid mothers who have a non-manual occupation and who stayed in full time education longest reporting the highest rates of colic in their infants
- Dietary change should not be the first intervention for colicky babies

The study deals only with what is reported and so no comment can be made about whether the infants labelled as colicky actually cried more than other infants. This does not represent a problem of interpretation for the application of the findings to clinical practice since health professionals usually have to advise on the basis of how infants are reported to behave. This may, however, have introduced imprecision into the ascertainment of colic at 1 month, which would have the effect of underestimating differences between colicky and non-colicky infants.

Interpretation of findings

A unifying interpretation of the main findings of this study is that the reporting of infantile colic by mothers is largely a social phenomenon. Mother's age and parity were the most important factors influencing the reporting of infantile colic, followed by the additional effect of socioeconomic factors. These findings are consistent with the hypothesis that parent-child interaction may be important in determining the reporting of colic. This interaction is influenced by a range of other factors including parents' expectations and interpretations of their infant's behaviour and the availability and use of coping mechanisms. The infant's sex did not influence the rate of reporting of colic, and this may argue against there being a profound psychosocial bias acting on the mother, since parents respond differently to male infants.[18] The design of this study, with quantitative information collected by questionnaire, means that it is not possible to examine the underlying causes of the reporting behaviour in any depth.

Babies who were reported to be colicky were also reported to sleep less and have much longer periods of wakefulness than non-colicky babies. As the amounts of sleep and wakefulness were not objectively measured, this finding could be subject to bias since mothers of colicky infants may be more likely to perceive that their infants do not sleep, or mothers of infants that do not sleep to perceive that their infant is colicky. However, babies who are reported to be colicky may indeed sleep less and may behave differently in other ways from non-colicky babies.

Mothers of breast fed infants reported colic at a higher rate than did mothers of bottle fed infants, but this effect was greatly reduced after adjustment for potential confounders in the logistic regression analysis. The confounder which had the greatest effect in reducing the independent effect of breast feeding was social class (measured by the proxies maternal education or occupation). As such social factors are poorly measured, their effects may be underestimated,[19] so the independent effect of breast feeding after adjustment may be the result of underadjustment for socioeconomic factors. If the type of milk feed influences the occurrence of infantile colic, its effect is either very small or takes place in a small minority of infants. Formula milk is most frequently derived from cows' milk. Since bottle feeding was associated with a lower odds of reporting colic than breast feeding, allergy to cows' milk protein is unlikely to account for the majority of reported colic.

The differences between colicky and non-colicky infants in birth weight and weight gain were statistically significant but were reduced by adjusting for confounders and in any case would not be regarded as clinically important. They were in the direction of colicky infants thriving.

At a population level, dietary factors contribute little to mother's reporting of infantile colic, and dietary change should not be the primary intervention. Future qualitative research is indicated to explore the social factors which influence the reporting of colic and interventional research should focus on both practical and psychosocial aspects in examining ways to alleviate the impact of persistent infant crying on families.

We thank Dr Elizabeth Taylor (University Department of Paediatrics, Sheffield) and Dr Robert Carpenter (London School of Hygiene and Tropical Medicine) for allowing access to the records of the Sheffield child development, and the health visitors of Sheffield, who collected the data.

1 Lester BM, Boukydis CFZ, Garcia-Coll CT, Hole WT. Colic for developmentalists. *Infant Mental Health J* 1990;11:321-33.
2 St James-Roberts I. Persistent crying in infancy. *Arch Dis Child* 1991;66:653-5.
3 Hide DW, Guyer BM. Prevalence of infantile colic. *Arch Dis Child* 1982;57:559-60.
4 Sloman J, Bellinger DC, Krentzel CP. Infantile colic and transient developmental lag in the first year of life. *Child Psychiatry Hum Dev* 1990;21:25-36.
5 Shaw C. A comparison of the patterns of mother-baby interaction for a group of crying, irritable babies and a group of more amenable babies. *Child Care Health Dev* 1977;3:1-12.
6 Keefe MR, Kotzer AM, Froese-Fretz A, Curtin M. A longitudinal comparison of irritable and nonirritable infants. *Nurs Res* 1996;45:4-9.
7 Barr R. The "colic" enigma: prolonged episodes of a normal predisposition to cry. *Infant Mental Health Journal* 1990;11:340-8.
8 Stacey TE. *National R&D programme of research in mother and child health.* London: NHS Executive, South Thames Regional Health Authority, 1995.
9 Carpenter RG, Gardner A, McWeeny PM, Emery JL. Multistage scoring system for identifying infants at risk of unexpected death. *Arch Dis Child* 1977;52:606-12.
10 Carpenter RG, Gardner A, Jepson M, Taylor EM, Salvin A, Sunderland R, *et al.* Prevention of unexpected infant death. Evaluation of the first seven years of the Sheffield intervention programme. *Lancet* 1983;i:723-7.
11 Rautava P, Helenius H, Lehtonen L. Psychosocial predisposing factors for infantile colic. *BMJ* 1993;307:600-4.
12 SAS. *Language and procedures.* Cary, NC: SAS Institute, 1990.
13 Altman DG. *Practical statistics for medical research.* London: Chapman and Hall, 1991.
14 Francis B, Green M, Payne C. *The GLIM system, release 4 manual.* Oxford: Clarendon Press, 1993.
15 Miller AR, Barr RG. Infantile colic. Is it a gut issue? *Pediatr Clin North Am* 1991;38:1407-23.
16 Stahlberg M-R. Infantile colic: occurrence and risk factors. *Eur J Pediatr* 1984;143:108-11.
17 Wessel MA, Cobb JC, Jackson EB, Harris G, Detwiler AC. Paroxysmal fussing in infancy, sometimes called "colic." *Pediatrics* 1954;14:421-34.
18 Lewis M. State as an infant-environment interaction: An analysis of mother-infant interaction as a function of sex. *Merril-Palmer Quarterly* 1972;18:95-121.
19 Phillips A, Davey Smith G. Bias in relative odds estimation owing to imprecise measurement of correlated exposures. *Stat Med* 1992;11: 935-61.

(Accepted 11 March 1997)