Papers

Randomised controlled trial of graded exercise in patients with the chronic fatigue syndrome

Kathy Y Fulcher, Peter D White

Abstract

Objective: To test the efficacy of a graded aerobic exercise programme in the chronic fatigue syndrome.
Design: Randomised controlled trial with control treatment crossover after the first follow up examination.
Setting: Chronic fatigue clinic in a general hospital department of psychiatry.
Subjects: 66 patients with the chronic fatigue syndrome who had neither a psychiatric disorder nor appreciable sleep disturbance.
Interventions: Random allocation to 12 weeks of either graded aerobic exercise or flexibility exercises and relaxation therapy. Patients who completed the flexibility programme were invited to cross over to the exercise programme afterwards.
Main outcome measure: The self rated clinical global impression change score, "very much better" or "much better" being considered as clinically important.
Results: Four patients receiving exercise and three receiving flexibility treatment dropped out before completion. 16 of 29 patients rated themselves as better after completing exercise treatment compared with eight of 30 patients who completed flexibility treatment. Analysis by intention to treat gave similar results (17/33 *v* 9/33 patients better). Fatigue, functional capacity, and fitness were significantly better after exercise than after flexibility treatment. 12 of 22 patients who crossed over to exercise after flexibility treatment rated themselves as better after completing exercise treatment. 32 of 47 patients rated themselves as better three months after completing supervised exercise treatment. 35 of 47 patients rated themselves as better one year after completing supervised exercise treatment.
Conclusion: These findings support the use of appropriately prescribed graded aerobic exercise in the management of patients with the chronic fatigue syndrome.

Introduction

Patients with the chronic fatigue syndrome perceive greater fatigue than healthy controls taking the same exercise.[1 2] This may be caused by physical deconditioning,[1] or sleep deprivation,[3] or psychological distress,[4] or a combination of the three. The syndrome seems to be independent of muscle strength and fatigability, which are normal in subjects with either the chronic fatigue syndrome[5 6] or the postinfectious fatigue syndrome.[7] Physical deconditioning, in turn, may be caused by reduced physical activity,[8] which may lead to adverse physical and psychological effects.[8 9]

Exercise provides physical and psychological benefits.[9 10] Graded aerobic exercise programmes can reduce incapacity and symptoms in many chronic and painful conditions, such as the post-polio syndrome,[11] chronic back pain,[12] and depressive illness.[10] Fitness training improved both aerobic capacity and myalgia more than flexibility exercises in patients with the "fibrositis/fibromyalgia syndrome,"[13] a condition which overlaps with the chronic fatigue syndrome.[14] A similar training programme improved symptoms and physiological findings in an open study of patients with the "effort syndrome."[15] There has been no randomised controlled trial of graded exercise treatment in patients with the chronic fatigue syndrome.[9]

We compared the physiological, symptomatic, and functional changes associated with a 12 week programme of either graded aerobic exercise or flexibility and relaxation therapy.

Patients and methods

Patients met the Oxford criteria for the chronic fatigue syndrome.[16] In addition, by using the structured clinical interview for the DSM-III-R (*Diagnostic and Statistical Manual of Mental Disorders,* third edition, revised)[17] we excluded patients who also had a current psychiatric disorder or symptomatic insomnia because of the separate effects of these conditions on fatigue.[3 4] We did not exclude patients with comorbid simple phobias.[18] Physical screening and investigations were carried out or, when appropriate, full recent records were obtained from the referring doctor to ensure other disorders had been excluded.[16]

In a previous study of exercise therapy in a related disorder half of the subjects considered themselves moderately improved by the treatment compared with 10% of controls receiving flexibility training.[13] By assuming similar treatment responses with $\alpha = 0.05$ and a power of 0.90 we calculated that 30 subjects would be required in each treatment group.[19] We recruited 66 patients to allow for 10% drop out.

In all, 167 outpatients were screened for the study. All had been referred to a chronic fatigue clinic in a

See editorial by Marcovitch

National Sports Medicine Institute, St Bartholomew's and the Royal London Medical School, London EC1M 6BQ
Kathy Y Fulcher, *laboratory director*
Department of Psychological Medicine, St Bartholomew's and the Royal London Medical School, London EC1A 7BE
Peter D White, *senior lecturer*
Correspondence to: Dr White.

BMJ 1997;314:1647–52

general hospital department of psychiatry. All patients with the chronic fatigue syndrome who also had a psychiatric disorder or insomnia were offered treatment for their comorbid disorder. If treatment was successful but the patients still met criteria for the chronic fatigue syndrome they were recruited into the trial. Ninety six patients did not meet criteria for the trial. Of these patients, 74 who had a current psychiatric disorder and three who had insomnia alone usually had their fatigue alleviated by treatment for these conditions. Nine patients had organic causes for their fatigue. A further five patients were not incapacitated enough, and five others were too incapacitated to attend for outpatient treatment. Only five of 71 (7%) eligible patients refused the trial.

Sixty six patients gave valid, informed consent and entered the trial. Of these, 27 (41%) had successfully been treated for a comorbid disorder beforehand but still met criteria for the chronic fatigue syndrome. Ethical approval was obtained from the district research ethics committee.

Assessment of symptoms

The hospital anxiety and depression scale measured anxiety and depression.[20] Fatigue was measured with self rated visual analogue scales measuring physical, mental, and total fatigue[18] and a self rated 14 item questionnaire.[21] The Pittsburgh sleep quality index measured quality and amount of sleep.[22] The 36 item short form health questionnaire (SF-36) was used specifically to measure physical functional capacity.[23]

Physiological assessments

The maximal voluntary contraction of the quadriceps muscle in the dominant leg was measured, with percutaneous twitch interpolation to ensure maximal activation.[2 24] A treadmill walking test was carried out at a constant 5 km/h, the slope being increased every two minutes. All patients were encouraged to continue the test to their maximum. Expired air was analysed continuously for percentages of oxygen and carbon dioxide and minute ventilation to give the peak oxygen consumption.

Patients rated their perceived exertion, using the unmodified Borg scale, in the last 30 seconds of each treadmill stage.[25] Thumb prick capillary lactate concentrations were measured at rest, at a perceived exertion score of 14 (between "somewhat hard" and "hard"), and three minutes after the test. The percentage of age predicted maximum heart rate reached was calculated from the formula $210 - (\text{age} \times 0.65)$.[26]

Treatments

All 66 patients were randomly allocated to either the exercise programme or the flexibility and relaxation programme, 33 patients being included in each group. Randomisation was achieved blindly to the psychiatrist and independently of the exercise physiologist by placing the letter E or F in 66 separate blank envelopes. These were then arranged in random order determined by random number tables and opened by an independent administrator after baseline tests as each new patient entered the study.

Exercise treatment

Patients attended weekly for 12 weeks of supervised treatment and the next week's exercise prescription. All laboratory sessions were supervised by an exercise physiologist using basic principles of exercise prescription,[27] which were adapted for the patient's current capacity. Home exercise was prescribed on at least five days a week, with initial sessions lasting between five and 15 minutes at an intensity of 40% of peak oxygen consumption (roughly 50% of the maximum recorded heart rate). The daily exercise prescription was increased by one or two minutes (negotiated with the patient each week) up to a maximum of 30 minutes. The intensity of exercise was then increased to a maximum of 60% of peak oxygen consumption. Patients were given ambulatory heart rate monitors to ensure that they reached but did not exceed target heart rates. The main exercise was walking, but patients were encouraged to take other modes of exercise, such as cycling and swimming. Patients were advised not to exceed prescribed exercise during a good phase. If patients complained of increased fatigue they were advised to continue at the same level of exercise for an extra week and increase when the fatigue had lessened.

Flexibility treatment

Patients attended the laboratory weekly for 12 weeks of flexibility and relaxation sessions provided by the same exercise physiologists. Each patient was taught a stretching routine and relaxation techniques. Patients were encouraged to start with sessions of 10 minutes, increasing to 30 minutes a day five days a week as more stretching exercises were added. They were specifically told to avoid doing any extra physical activities.

Patients from each group attended the laboratory at different times to avoid communication between groups. Patients kept a weekly activity diary, recording the type, duration, and response to exercise or stretching, which determined the next week's prescription.

Outcome measures

The main outcome measure was the self rated clinical global impression change score, which is a validated measure of overall change compared with study onset, with seven possible scores from "very much worse" (score 7) to "very much better" (score 1).[28] Secondary outcome measures included assessments of strength and fitness, symptoms, and functional capacity, as described above.

Follow up

After 12 weeks all patients were reassessed as at baseline except that the mid-exercise blood lactate concentration was measured at the same treadmill stage as before. Patients were blind to all measures at all stages. At any time two exercise physiologists participated in both testing and treatment. During the follow up week patients also attended a psychiatrist (who was blind to treatment) for a structured clinical interview for the DSM-III-R and to record their self rated clinical global impression change score. Seven subjects (four in the exercise group, three in the flexibility group) did not attend the psychiatrist at the correct time despite completing treatment; these patients assessed their clinical global impression score retrospectively and returned it by post.

Table 1 Self rated clinical global impression change scores after completing treatment

Clinical global impression change score	No (%) in exercise group (n=29)	No (%) in flexibility group (n=30)
1 (Very much better)	9 (31)	2 (7)
2 (Much better)	7 (24)	6 (20)
3 (A little better)	11 (38)	18 (60)
4 (No change)	1 (3)	3 (10)
5 (A little worse)	1 (3)	0
6 (Much worse)	0	1 (3)
7 (Very much worse)	0	0

After the follow up reassessment patients in the flexibility group could cross over to exercise treatment. Treatment and follow up were the same as for the original exercise group.

All patients were instructed to continue with regular exercise and could remain in contact with the laboratory, mainly by phone. Three months after stopping supervised exercise patients were reassessed. Roughly one year after supervised treatment stopped a letter was sent to all patients asking about their activities and asking them to record their self rated clinical global impression change score.

Statistical analysis

The clinical global impression change score was analysed categorically, a score of 1 or 2 ("very much better" or "much better") being considered clinically important versus scores of 3 to 7 ("a little better" to "very much worse"). We compared the proportions of patients rating themselves as clinically improved among those who completed treatment as well as by intention to treat analysis by means of a χ^2 test with a continuity correction. We completed follow up assessments on four of the seven patients who dropped out of treatment and included these data in the intention to treat analysis. Patients with missing data were counted as non-improvers. We compared all subsidiary outcome variables by Student's t or Mann-Whitney tests. The effects of the two treatments on the submaximal responses to exercise were compared by examining the group means of heart rate and perceived exertion score during the middle third (6-12 minutes inclusive) of the group maximum treadmill tests. At the three month and one year follow ups we reported the proportions of patients feeling better both after completing treatment and by intention to treat analysis.

Results

Baseline data

The 66 patients had a mean age of 37.2 (SD 10.7) years and a mean body mass index of 23.8 (4.6). Forty nine (74%) were women. The median duration of illness was 2.7 (range 0.6-19.0) years. Twenty patients were taking full dose antidepressants; 10 were taking low dose tricyclic antidepressants as hypnotics. All patients were told to continue their medication unchanged. Two thirds of patients (n=44) blamed viruses for their illness. There were no significant differences in the proportions of both groups taking antidepressants or blaming a virus for their illness.

Treatment results

Table 1 shows the main outcome of treatment. Sixteen of the 29 patients who completed exercise treatment rated themselves as "much" or "very much" better compared with eight of the 30 patients in the flexibility group ($\chi^2=3.85$, df=1; P=0.05). Only two patients (one in each group) rated themselves worse after treatment; both had developed a major depressive illness. Four patients in the exercise group and three in the flexibility group dropped out of treatment. Only two patients (one from each group) said they dropped out because the treatment made them worse. Analysis by intention to treat showed that 17 of 33 patients improved with exercise and nine of 33 improved with flexibility treatment ($\chi^2=4.06$, df=1; P=0.04).

The median peak oxygen consumption was significantly greater after exercise than after flexibility treatment (13% v 6% increase; table 2). The median increase in isometric strength was 26% in the exercise group and 15% in the flexibility group (Mann-Whitney U test=351; P=0.20). The mean heart rate during submaximal treadmill testing was 143 (SD 13) beats/min after exercise treatment versus 150 (13) beats/min after flexibility treatment ($t=2.06$, df=56; P=0.04) (fig 1). The mean submaximal perceived exertion score was 14.5 (3.4) with exercise compared with 16.2 (2.8) with flexibility treatment ($t=2.13$, df=57; P=0.04) (fig 2). None of these peak or submaximal measures was significantly different before treatment. Those patients in the exercise group who rated themselves as better had no significantly greater improvement in either peak oxygen consumption ($t=0.71$,

Table 2 Physiological variables at baseline and after completing treatment

Variable	Baseline		After 12 weeks of treatment		P value
	Exercise group (n=33)	Flexibility group (n=33)	Exercise group (n=29)	Flexibility group (n=30)	
Median (interquartile range) peak oxygen consumption with exercise (ml/kg/min)	31.8 (26.8-36.8)	28.2 (23.4-33.1)	35.8 (30.8-40.7)	29.8 (24.7-34.9)	0.03
Median (interquartile range) maximum ventilation (l/min)	71.2 (52.6-89.7)	69.7 (62.0-77.5)	88.6 (66.9-107.3)	76.1 (54.0-98.2)	0.04
Mean (SD) maximum heart rate (beats/min)	170 (18)	171 (19)	174 (17)	178 (14)	0.24
Median (interquartile range) percentage of predicted maximum heart rate	91 (85-98)	93 (89-97)	95 (91-99)	95 (90-100)	0.34
Mean (SD) recovery heart rate three minutes after test (beats/min)	109 (18)	111 (17)	111 (16)	115 (15)	0.68
Mean (SD) test duration (min)	10.5 (3.7)	9.5 (3.6)	12.4 (3.5)	11.0 (3.3)	0.08
Median (interquartile range) submaximal blood lactate (mmol/l)	2.6 (1.4-3.8)	2.3 (1.4-3.4)	2.0 (1.2-2.8)	2.5 (1.6-3.4)	0.91
Mean (SD) post-test blood lactate (mmol/l)	4.9 (1.9)	5.8 (2.5)	6.2 (2.5)	6.1 (2.5)	0.95
Mean (SD) maximal quadriceps voluntary contraction (with twitch interpolation) (N)	339 (144)	340 (105)	430 (182)	378 (100)	0.18

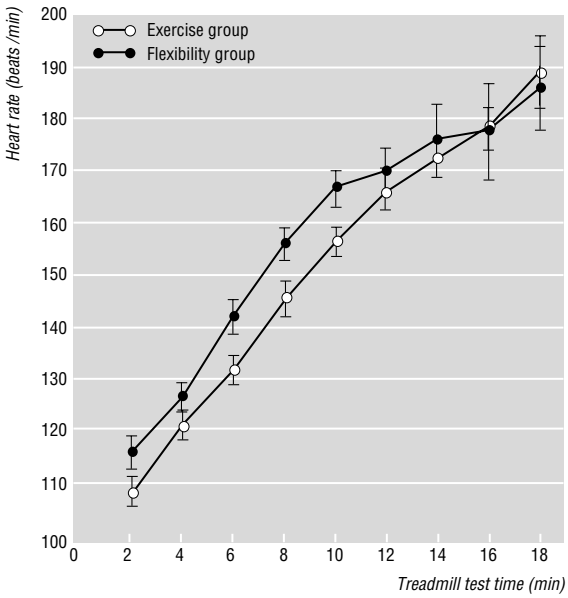

Fig 1 Mean heart rates during treadmill testing after treatment. Bars are SEM

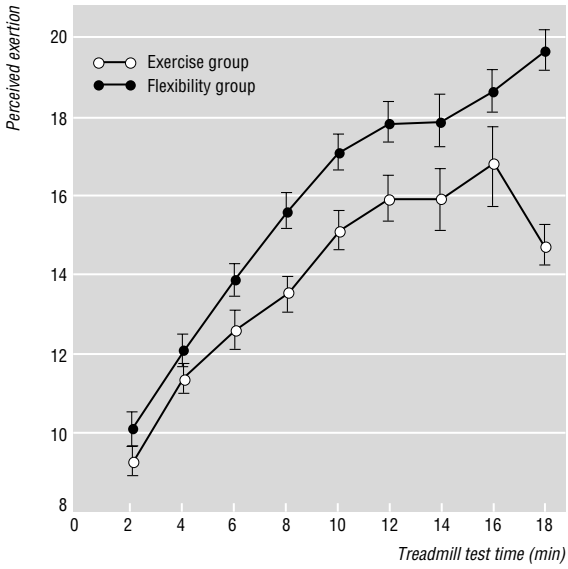

Fig 2 Mean perceived exertion scores during treadmill testing after treatment. Bars are SEM

df = 26; P = 0.48) or strength (Mann-Whitney U test = 71; P = 0.24) than the rest.

Both total and physical fatigue, total SF-36 score, physical function, and general health were significantly better after exercise than after flexibility treatment, though none of these measures had returned to normal (table 3).

Crossover subjects

Seven patients who completed the flexibility programme chose not to proceed to the exercise treatment, usually because of domestic commitments. Twenty three patients went on to exercise treatment immediately after the flexibility treatment. Twelve of 22 patients who completed the exercise programme rated themselves as better at follow up. One subject had a foot operation for an unrelated condition and dropped out. Significant improvements occurred in peak oxygen consumption ($P < 0.0001$) and physical function ($P = 0.002$) compared with baseline, which had not significantly improved after flexibility alone.

Three month follow up after exercise

Forty seven of the original 66 patients were reassessed three months after stopping supervised exercise treatment. Of the remaining 19 patients, seven dropped out of the first three months of treatment, seven did not cross from flexibility to exercise treatment, and five could not attend because of moving home or an injury unrelated to exercise. Thirty two of those who attended follow up after completing exercise treatment rated themselves as better. By intention to treat analysis 32 of 56 patients were better after exercise.

Physiological improvements recorded at first follow up were maintained or exceeded in patients who attended three months after completing exercise treatment. In particular, the mean maximal voluntary contraction force of the quadriceps muscle had increased from 339 (SD 126) Newtons (N) at baseline to 436 (161) N by six months ($P < 0.001$).

One year follow up

Fifty two (79%) patients returned questionnaires one year after stopping their supervised treatment. Thirty five of 47 (74%) patients who completed exercise treatment rated themselves as better (35 of 56 (63%) patients by intention to treat analysis). Of the seven

Table 3 Symptomatic and functional measures at baseline and after completing treatment

	Baseline		After 12 weeks of treatment		
Variable	**Exercise group**	**Flexibility group**	**Exercise group**	**Flexibility group**	**P value**
Mean (SD) Chalder fatigue score	28.9 (7.1)	30.5 (5.6)	20.5 (8.9)	27.4 (7.4)	0.004
Mean (SD) total fatigue score†	312 (50)	325 (45)	253 (48)	286 (67)	0.04
Mean (SD) physical fatigue score†	161 (27)	177 (16)	130 (28)	154 (34)	0.006
Mean (SD) mental fatigue score†	151 (26)	148 (34)	124 (31)	132 (39)	0.38
Mean (SD) SF-36 total score	341 (84)	336 (73)	478 (112)	420 (120)	0.05
Mean (SD) SF-36 physical function score	48.5 (22.1)	47 (18.7)	69 (18.5)	55 (21.8)	0.01
Median (interquartile range) SF-36 general health score	41 (22-48)	33 (21-40)	45 (36-66)	37 (25-52)	0.03
Median (interquartile range) depression score‡	5 (1.5-8.5)	5 (2.5-7.5)	5.5 (2.9-8.1)	4 (0.6-7.4)	0.92
Median (interquartile range) anxiety score‡	5.5 (1.5-8.5)	4 (0.5-7.5)	5.5 (3.0-8.0)	7 (3.5-10.5)	0.46
Median (interquartile range) sleep total score§	7 (5.5-8.5)	6 (4.5-7.5)	5 (3.5-6.5)	6 (4.1-7.9)	0.49

Normal or usual scores are 14 for Chalder questionnaire, 200 for total fatigue score (visual analogue scale), and 100 for physical and mental fatigue scores (visual analogue scales). 100 is maximum (full capacity) SF-36 score for physical and general health function. Score less than 8 on hospital anxiety and depression scale is considered non-pathological. Pittsburgh sleep quality index score less than 6 is considered non-pathological.
†Visual analogue scale.
‡Hospital anxiety and depression scale.
§Pittsburgh sleep quality index.

patients who had flexibility treatment alone, five replied, of whom two felt better (one of whom had done her own exercise programme). Thirty one (66%) patients who completed exercise treatment were working or studying at least part time compared with 26 (39%) of all 66 patients before treatment (95% confidence interval of difference 9% to 44%). Thirty two of 47 (68%) of these patients had continued exercising and considered themselves regularly active compared with 36 of 65 (55%) patients who said they were regularly active before becoming ill (no significant difference).

Discussion

Graded exercise treatment was more effective than relaxation and stretching exercises, suggesting that the amount of therapists' attention was not responsible for the difference in outcome. Analysis by intention to treat gave similar results, with low drop out rates and minimal adverse effects. Only five patients declined the study and five were too ill to attend as outpatients. The same result in patients who crossed over supports the original finding. The subsidiary outcome measures of fatigue and functional capacity confirmed the greater improvement with exercise. Improvement was maintained or exceeded at both three and 12 months of follow up. Almost three quarters of patients followed up felt better and had an accompanying return to premorbid levels of physical activity, and an increased proportion had returned to work or study. In a similar sample only 2% of subjects reported spontaneous resolution of fatigue by 18 months of follow up.[29] Hence it is unlikely that spontaneous improvement would have occurred without exercise in our series. Nevertheless, because of the crossover design of our study we cannot state with certainty that exercise was responsible for the continued improvement after the first follow up examination.

The 13% increase in aerobic capacity was consistent with an increase of between 5% and 10% found in healthy but sedentary people performing a similar training programme.[30] Surprisingly, both groups significantly increased their strength with treatment despite previous studies finding no significant difference in the strength of sufferers when compared with healthy controls.[5-7] These improvements in strength may have been due to the increased mobility that occurred with both treatments. However, patients given flexibility treatment did not significantly improve their aerobic capacity until they crossed over to exercise treatment.

Both physiological and perceptual improvements occurred at submaximal treadmill stages after exercise treatment compared with flexibility treatment, as reported in an open study.[15] This is clinically important, as submaximal activities such as walking are functionally more important than maximal activities such as running. The different changes in heart rate response to exercise were not related to antidepressants, as similar proportions of patients from both groups were taking these drugs.

The limitations of our study include the exclusion of 44% (74/167) of possible subjects because of their comorbid psychiatric disorders, though this was mitigated by the inclusion of 16% (27/167) of subjects who had their psychiatric disorder successfully treated before the study. Though bias may have been introduced by the lack of blindness in the exercise physiologists at reassessment, care was taken to standardise the encouragement given; submaximal results would not have been affected, and the main outcome measure was self rated to avoid this bias.

We can only speculate whether it was important to treat or exclude patients with a psychiatric disorder or sleep disturbance. The only two patients to feel worse after completing treatment had developed depressive illnesses. Though exercise may help depressive illness,[10] it improved neither mood nor sleep in this or a previous study.[13] The present evidence suggests that these factors should be treated before starting an exercise programme.[3 4]

The only other treatment of the chronic fatigue syndrome to show promise is cognitive behaviour therapy, which improves functional capacity and symptoms more than both standard medical care and relaxation therapy.[31 32] At first glance exercise treatment seems to work more quickly, as 52% of subjects had improved (by intention to treat analysis) after three months compared with 27% with cognitive behaviour therapy after five months.[31] However, outcomes at final follow up were similar. The treatments cannot fairly be compared, as one cognitive behaviour therapy study recruited subjects with greater perceived physical incapacity[32] and both cognitive behaviour therapy studies treated subjects with concurrent psychiatric disorders. Whether cognitive behaviour therapy is equally effective in the absence of psychiatric disorders is uncertain.[31 32]

Key messages

- Graded exercise treatment is more effective than relaxation and flexibility treatment for patients with the chronic fatigue syndrome who do not have a psychiatric disorder or sleep disturbance
- Overall improvement is accompanied by improvements in fatigue and physical function but seems independent of the improved strength and peak aerobic capacity produced by exercise
- In this survey few patients refused or dropped out of exercise treatment and only one patient claimed to be worse after completing it
- Patients show sustained benefit one year after graded exercise treatment

We are grateful for the help and advice of Ms S Gibbs, Ms J Thomas, Dr D Tunstall-Pedoe, Dr K J Cleary, and Professor C Williams.

Funding: The study was supported by a grant from the Linbury Trust, a Sainsbury charitable trust.

Conflict of interest: None.

1 Riley MS, O'Brien CJ, McCluskey DR, Bell NP, Nicholls DP. Aerobic work capacity in patients with chronic fatigue syndrome. *BMJ* 1990;301:953-6.
2 Gibson H, Carroll N, Clague JE, Edwards RHT. Exercise performance and fatigability in patients with CFS. *J Neurol Neurosurg Psychiatry* 1993;56:993-8.
3 Myles WS. Sleep deprivation, physical fatigue, and the perception of exercise intensity. *Med Sci Sports Exerc* 1985;17:580-4.
4 Morgan WP. Psychological components of effort sense. *Med Sci Sports Exerc* 1994;26:1071-7.

5 Stokes MJ, Cooper RG, Edwards RH. Normal muscle strength and fatigability in patients with effort syndrome. *BMJ* 1988;297:1014-7.
6 Lloyd A, Gandevia S, Hale J. Muscle performance, voluntary activation, twitch properties and perceived effort in normal subjects and patients with the chronic fatigue syndrome. *Brain* 1991;114:85-98.
7 Rutherford OM, White PD. Human quadriceps strength and fatigability in patients with post viral fatigue. *J Neurol Neurosurg Psychiatry* 1991;54:961-4.
8 Saltin B, Blomquist G, Mitchell J, Johnson R, Wildenthal K, Chapman C. Response to exercise after bedrest and after training. *Circulation* 1968;38:1-78.
9 McCully K, Sisto S, Natelson B. Use of exercise for treatment of chronic fatigue syndrome. *Sports Med* 1996;21:35-48.
10 Weyerer S, Kupfer B. Physical exercise and psychological health. *Sports Med* 1994;17:108-16.
11 Feldman RM, Soskolne CL. The use of nonfatiguing strengthening exercises in post-polio syndrome. *Birth Defects* 1987;23:335-41.
12 Frost H, Klaber Moffett JA, Moser JS, Fairbank JCT. Randomised controlled trial for evaluation of fitness programme for patients with chronic low back pain. *BMJ* 1995;310:151-4.
13 McCain GA, Bell DA, Mai FM, Halliday PD. A controlled study of the effects of a supervised cardiovascular fitness training program on the manifestations of primary fibromyalgia. *Arthritis Rheum* 1988;31: 1135-41.
14 Buchwald D, Garrity D. Comparison of patients with chronic fatigue syndrome, fibromyalgia, and multiple chemical sensitivities. *Arch Intern Med* 1994;154:2049-53.
15 Newham D, Edwards RH. Effort syndromes. *Physiotherapy* 1979;65:52-6.
16 Sharpe MC, Archard LC, Banatvala JE, Borysiewicz LK, Clare AW, David A, et al. A report—chronic fatigue syndrome: guidelines for research. *J R Soc Med* 1991;84:118-21.
17 Spitzer RL, Williams JBW, Gibbon M, First MB. *Structured clinical interview for DSM-III-R: patient edition.* Washington, DC: American Psychiatric Press, 1990.
18 White PD, Grover SA, Kangro HO, Thomas JM, Amess J, Clare AW. The validity and reliability of a fatigue syndrome that follows glandular fever. *Psychol Med* 1995;25:917-24.
19 Fleiss JL. *Statistical methods for rates and proportions.* 2nd ed. Chichester: J Wiley & Sons, 1981.
20 Zigmond A, Snaith R. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;87:361-70.
21 Chalder T, Berelowitz G, Pawlikowska T, Watts L, Wessely S, Wright D, et al. Development of a fatigue scale. *J Psychosom Res* 1993;37:147-53.
22 Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatr Res* 1989;28:193-213.
23 Ware JE, Sherbourne CD. The MOS 36-item short form health survey (SF36). *Med Care* 1992;30:473-83.
24 Rutherford OM, Newham D. Clinical and experimental application of the percutaneous twitch superimposition technique for the study of human muscle weakness. *J Neurol Neurosurg Psychiatry* 1986;49:1288-91.
25 Borg G. Perceived exertion as an indicator of somatic stress. *Scand J Rehabil Med* 1970;2:92-8.
26 Jones NL. *Clinical exercise testing.* Philadelphia: Saunders, 1988.
27 American College of Sports Medicine. *Guidelines for exercise testing and prescription.* 4th ed. Philadelphia: Lea & Febiger, 1991.
28 Guy W. *ECDEU assessment manual for psychopharmacology.* Rockville, MD: National Institute of Mental Health, 1976:218-22.
29 Bombardier CH, Buchwald D. Outcome and prognosis of patients with chronic fatigue vs chronic fatigue syndrome. *Arch Intern Med* 1995;155:2105-10.
30 Blomqvist C, Saltin B. Cardiovascular adaptations to physical training. *Ann Rev Physiol* 1983;45:169-79.
31 Sharpe MC, Hawton K, Simkin S, Surawy C, Hackmann A, Klimes A, et al. Cognitive behaviour therapy for the chronic fatigue syndrome: a randomised controlled trial. *BMJ* 1996;312:22-6.
32 Deale A, Chalder T, Marks I, Wessely S. Cognitive behavior therapy for chronic fatigue syndrome: a randomised controlled trial. *Am J Psychiatry* 1997;154:408-14.

(Accepted 8 April 1997)

Discontinuation of cervical spine immobilisation in unconscious patients with trauma in intensive care units—telephone survey of practice in South and West region

K J Gupta, M Clancy

Department of Anaesthesia, Frenchay Hospital, Bristol BS16 1LE
K J Gupta, *specialist registrar*

Department of Accident and Emergency, Bristol Royal Infirmary, Bristol BS2 8HW
M Clancy, *consultant*

Correspondence to: Dr K J Gupta, Department of Anaesthesia, Royal United Hospital, Combe Park, Bath BA1 3NG.

BMJ 1997;314:1652–5

Abstract

Objective: To study how the cervical spine is assessed before discontinuation of cervical spine immobilisation in unconscious trauma patients in intensive care units.
Design: Telephone interview of consultants responsible for adult intensive care units.
Setting: All 25 intensive care units in the South and West region that admit victims of major trauma.
Main outcome measures: The clinical and radiological basis on which the decision is made to stop cervical spine immobilisation in unconscious patients with trauma.
Results: In 19 units cervical spine immobilisation was stopped in unconscious patients on the basis of radiology alone, and six units combined radiology with clinical examination after the patient had regained consciousness. Sixteen units relied on a normal lateral radiological view of the cervical spine alone, five required a normal lateral and anteroposterior view, and four required a normal lateral, anteroposterior, and open mouth peg view.
Conclusions: There are inconsistencies in the clinical and radiological approach to assessing the cervical spine in unconscious patients with trauma before the removal of immobilisation precautions. There is an overreliance on the lateral cervical spine view alone, which has been shown to be insensitive in this setting.

Introduction

The reported incidence of cervical spine injuries in victims of major trauma varies from 2.3% to 12%.[1-3] Victims of trauma therefore have their cervical spine immobilised at the earliest possible opportunity. Those patients with a severe head injury (Glasgow coma scale <8) are usually rapidly anaesthetised and intubated as part of their emergency care and are subsequently admitted to an intensive care unit. The cervical spine will normally remain immobilised during these procedures until injury to the neck can be ruled out.

In the advanced trauma life support course for physicians the American College of Surgeons' committee on trauma states that "Patients...should be presumed to have an unstable cervical spine injury and the neck should be immobilized until all aspects of the cervical spine have been adequately studied and an injury excluded." It also states that discontinuation of spinal immobilisation "usually occurs when no roentgenographic abnormality has been documented, and no symptoms or signs relating to the spine or cord exist."[4] While this has become the accepted

standard in conscious patients it is less easily applied if the patient is unconscious. The assessment of the cervical spine before removal of immobilisation precautions in unconscious trauma patients therefore remains controversial.

Methods

KJG conducted a telephone survey of all of the general intensive care units dealing with major adult trauma in a single NHS region (South and West) over a period of two weeks. These units were identified from *The Directory of Emergency and Special Care Units*.[5] Units were excluded only if they were in hospitals without an accident and emergency department from which to admit victims of trauma. The consultant in charge of the intensive care unit on the day of the call was interviewed, or the clinical director of the unit if the consultant was not available. If neither were available then the call was made on another day when another consultant was covering the unit.

The participating consultants were given the same case scenario of an unconscious and intubated trauma victim with a head injury being admitted to their intensive care unit. They were told that the patient had no specific clinical evidence of a cervical spine injury but that the patient's cervical spine was immobilised with a semirigid cervical collar.

They were asked what clinical and radiological investigations were normally required on their unit before such immobilisation was discontinued. Each consultant was then allowed to answer without further prompting or interruption. If specific information required for the survey was not obtained in the first answer a prepared structured question pertaining to that specific aspect was used.

The specific information sought during the interview was, firstly, whether the cervical spine immobilisation precautions were normally discontinued before or after the patient had regained consciousness; secondly, which radiological views were required as a minimum, assuming each view was anatomically complete and of good quality; thirdly, the specialty and grade of those normally interpreting the film; fourthly, whether the hospital or unit had an established policy or guidelines for this problem and, if so, whether the responses were based on this; fifthly, if the respondents had in their practice encountered morbidity from the prolonged use of a semirigid cervical collar; and, finally, if the respondents had encountered any patients with a missed cervical spine injury detected only after removal of immobilisation precautions.

Results

All 25 suitable intensive care units responded to the telephone survey. In each unit the information was provided by a consultant with a regular commitment to intensive care. Sixteen of the units were in hospitals whose accident and emergency departments saw over 30 000 patients a year, seven in hospitals seeing over 50 000 patients a year, and two in hospitals seeing more than 70 000 accident and emergency patients a year. Three of the hospitals contained neurosurgical centres, and one was a centre for spinal injuries.

Table 1 Minimum radiological requirements before removal of cervical spine immobilisation precautions according to consciousness of patients. Values are numbers of hospital units

Detail	Lateral only	Lateral and anteroposterior	Lateral, anteroposterior, and open mouth peg view
Unconscious patients	12	5	2
Patients who regained consciousness	4	0	2

Ten consultants responded according to guidelines relating to this problem already established in their hospital. These guidelines had been set in three cases by intensivists, three by neurosurgeons or orthopaedic surgeons, one by radiologists, and three by a multidisciplinary team. Fifteen units had no established guidelines or protocol in place.

The practice on 19 units was to remove all cervical spine precautions while the patient was still unconscious. Six would maintain cervical spine immobilisation until patients had regained consciousness. Table 1 shows the minimum radiological requirements accepted by units with these different criteria.

There was a large overlap in the specialty of the doctor commonly asked to interpret the radiographs. Five units regularly used a radiologist, and seven units regularly used either an orthopaedic surgeon or neurosurgeon. Ten units would use either specialist depending on which was available. On three of the units the intensive care consultants would regularly interpret the films. All the units used grades of senior registrar or above in the relevant specialty.

Eighteen of the consultants had, in their practice, encountered morbidity associated with the prolonged use of semirigid cervical collars on intensive care patients: 10 related to the occurrence of occipital or mental pressure sores, or both; four had experienced considerable difficulties during extubation or reintubation of patients in collars; and two volunteered that nursing staff had complained of general difficulties, including awkward mouth care and increased nursing requirements during "log roll" turns of patients.

Four of the consultants were aware of cases of missed injury of the cervical spine, discovered after removal of immobilisation precautions. Information on the outcome in these patients was not available.

Discussion

Methods

Telephone surveys are more effective than postal surveys in achieving complete participation as posted forms may not be received or may be mislaid or ignored. We were thus able to survey practice in all the hospitals in our region that may receive victims of major trauma. A criticism of telephone interviews is that the interaction they afford allows the interviewer inadvertently to bias the respondents. We minimised this by allowing each consultant to give a full and uninterrupted answer after an initial open ended question relating to a case scenario. Only if specific information required for the survey was not obtained from the first answer did interaction take place. This was in the form of specific structured questions.

In any telephone survey, assessment of the reliability of responses and their applicability to actual practice may be difficult. It is important to interview the person responsible for making the relevant decision.[6] We therefore interviewed the consultants responsible for the unit on the day of the interview. Respondents may also be influenced by personal opinion based on a perceived "ideal" practice. This is difficult to control, but the consultants were asked to describe the normal practice in their unit in an attempt to minimise the influence of personal bias, particularly in those units without established guidelines. It is also possible that consultants from different hospitals could influence each other's answers through discussion of this problem at local or regional meetings. We therefore carried out the survey over a brief period of time, hoping to achieve a "snap shot" of practice across the region.

Findings in the context of current knowledge

Sixteen of the 25 (64%) units were prepared to rely on a lateral plain radiograph as their sole cervical spine radiology. Twelve of these units would be prepared to remove cervical spine immobilisation on the basis of this view while the patient was still unconscious. Several studies have clearly shown that the cross table lateral view in isolation will miss about 15% of patients with a cervical spine fracture or dislocation, even if the films are anatomically complete, of good quality, and are read by an expert.[2 7 8] Between 37% and 63% of these missed injuries will be potentially unstable.[7 8] The problem of lateral views is further compounded when the injuries are missed because of inexpert interpretation or incomplete films, or both.[1-3 8] The commonest areas of missed injuries on the lateral view are the C1/2 and C7/T1 regions.[8 9] Between 7% and 25% of lateral views will not show the top of T1, despite the use of arm traction.[3 10 11]

Five of the 25 centres reported using three plain views to visualise the cervical spine: the lateral, anteroposterior, and open mouth odontoid peg view. Several studies have shown the combination of these views to have a sensitivity of 90-99% for the detection of cervical spine injuries.[2 7-9 11 12] Again, expert interpretation of the films is necessary to achieve these rates. In our survey 22 units were specifically asking a radiologist, a senior orthopaedic surgeon, or a neurosurgeon to interpret the radiographs.

None of the consultants said that their units were using computed tomography as a first line investigation, although several were using it as a second line investigation if plain radiography was inadequate. Computed tomography can miss horizontally orientated fractures,[12-14] although this risk can be reduced by the use of reconstructed sagittal images. The sensitivity of computed tomography when used alone varies from 80% to 92%.[11 13 14] If it is used in combination with three plain view films by being targeted to poorly visualised or suspicious areas on these films, however, the sensitivity for the detection of cervical spine injury approaches 100%.[9 11 14]

Six units would wait until the patients had regained consciousness before removing immobilisation precautions. We were unable to find any published evidence in unconscious patients to support or refute this practice. It seems sensible to wait for the added reassurance of normal results from a clinical examination in these patients. Several papers, however, have highlighted the poor reliability of such normal results in patients with spine injury who have an altered consciousness because of head injury, sedative drugs, or alcohol, or who have other distracting injuries.[8 15 16] Also the effectiveness of semirigid cervical collars in providing immobilisation when used alone is questionable[19 20] and, as implied by the anecdotal reports in this survey, longer term use of semirigid cervical collars can be associated with considerable morbidity.[17 18]

Key messages

- Immobilisation of the cervical spine is essential in victims of major trauma until the cervical spine is fully assessed for injury
- In the NHS region surveyed intensive care units show little consistency in how the cervical spine is assessed before the removal of immobilisation precautions
- Most practitioners would be prepared to accept radiography of the lateral cervical spine as the sole radiological assessment, although it has been shown to be an insensitive screening tool
- Computed tomography targeted at poorly visualised or suspicious areas in three plain radiographs (lateral, anteroposterior, and open mouth peg views) is a sensitive technique for the radiological detection of injury to the bony cervical spine

What is the incidence of the problem?

Four consultants knew of a patient in whom a cervical spine injury was diagnosed only after removal of immobilisation precautions. These reports are anecdotal, but we would not expect an intensive care unit to reproduce figures on a problem it would encounter comparatively infrequently. Anecdotes are often the only source of information on uncommon events and therefore these results are still relevant. Some reports claim that injuries of the cervical spine are missed or their diagnosis is delayed in 4.6% to 17% of victims of trauma with spine injury.[1 2 21] Delay in diagnosis of cervical spine injuries has been shown to effect mortality and morbidity, with up to 29% of patients suffering neurological deterioration as a result.[1 22] While some missed injuries may be the result of isolated ligamentous disruption, which may not be visible on plain radiography,[23] the incidence of this type of injury in this population is probably less than 1%.[17]

Conclusions

The approach to the common problem of deciding when to discontinue cervical spine immobilisation in unconscious patients with trauma in an intensive care unit varies widely in the South and West region. There is no consistency in the minimum radiological investigations required or whether the patient's consciousness should effect the decision. There is an

overdependence on the isolated use of lateral plain radiographs, which have been shown to be an insensitive screening tool. The detection rate of injury can be considerably improved by the use of three plain views (a lateral, anteroposterior, and open mouth peg view), particularly if poorly visualised or suspicious areas seen on these views are targeted with computed tomography.

Funding: No external funding.
Conflict of interest: None.

1 Davis JW, Phreaner DL, Hoyt DB, Mackersie RC. The etiology of missed cervical spine injuries. *J Trauma* 1993;**4**:342-6.
2 MacDonald RL, Schwartz ML, Mirich D, Sharkey PW, Nelson WR. Diagnosis of cervical spine injury in motor vehicle crash victims: how many x-rays are enough? *J Trauma* 1990;30:392-7.
3 Cohn SM, Lyle WG, Linden CH, Lancey RA. Exclusion of cervical spine injury: a prospective study. *J Trauma* 1991;31:570-4.
4 American College of Surgeons. *Advanced trauma life support course for physicians.* Chicago: Committee on Trauma, American College of Surgeons; 1993:31/194.
5 CMA Medical Data. *Directory of emergency and special care units 1996.* Cambridge: CMA Medical Data, Cambridge Research Laboratories, 1996.
6 Palmer JD, Wagstaff A, McKelvie G. Answers may have reflected perceived rather than actual management (letter). *BMJ* 1996;313:296.
7 Streitwieser DR, Knopp R, Wales LR, Williams JL, Tonnemacher K. Accuracy of standard radiographic views in detecting cervical spine fractures. *Ann Emerg Med* 1983;12:538-42.
8 Woodring JH, Lee C. Limitations of cervical radiography in the evaluation of acute cervical trauma. *J Trauma* 1993;34:32-9.
9 Borock EC, Gabram SGA, Jacobs LM, Murphy MA. A prospective analysis of a two year experience using computed tomography as an adjunct for cervical spine clearance. *J Trauma* 1991;31:1001-5.
10 Mirvis SE, Diaconis JN, Chirico PA, Reiner BI, Joslyn JN, Militello P. Protocol driven radiologic evaluation of suspected cervical spine injury: efficacy study. *Radiology* 1989;170:831-4.
11 Ross SE, Schwab CW, David ET, Delong WG, Born CT. Clearing the cervical spine: initial radiologic evaluation. *J Trauma* 1987;27:1055-9.
12 Acheson MB, Livingston RR, Richardson ML, Stimac GK. High resolution CT scanning in the evaluation of cervical spine fractures: comparison with plain film examinations. *Am J Roentgenol* 1987;148:1179- 85.
13 Schleehauf K, Ross SE, Civil ID, Schwab CW. Computed tomography in the initial evaluation of the cervical spine. *Ann Emerg Med* 1989;18:815-7.
14 Woodring JH, Lee CL. The role and limitations of computed tomographic scanning in the evaluation of cervical trauma. *J Trauma* 1992;33:698-708.
15 Hoffman JR, Schriger DL, Mower W, Luo JS, Zucker M. Low risk criteria for cervical spine radiography in blunt trauma: a prospective study. *Ann Emerg Med* 1992;21:1454-60.
16 Roberge RJ, Wears RC, Kelly M, Evans TC, Kenny M, Daffner RD, *et al.* Selective application of cervical spine radiography in alert victims of blunt trauma: a prospective study. *J Trauma* 1988;28:784-8.
17 Davis JW, Parks SN, Detlefs CL, Williams GG, Williams JL, Smith RW. Clearing the cervical spine in obtunded patients: the use of dynamic fluoroscopy. *J Trauma* 1995;39:435-8.
18 Plaiser B, Gabram SGA, Schwartz RJ, Jacobs LM. Prospective evaluation of craniofacial pressure in four different cervical orthoses. *J Trauma* 1994;37:714-9.
19 Grady MS, Howard MA, Jane JA. The use of the Philadelphia collar as an alternative to the Halo vest in patients with C-2, C-3 fractures. *Neurosurgery* 1986;18:151.
20 Rosen PB, McSwain NE, Arata M.. Comparison of two new immobilization collars. *Ann Emerg Med* 1992;21:1189 .
21 Gerrelts BD, Petersen EU, Mabry J, Petersen SR. Delayed diagnosis of cervical spine injuries. *J Trauma* 1991;31:1622-6.
22 Reid DC, Henderson R, Saboe L, Miller JDR. Etiology and clinical course of missed spine fractures. *J Trauma* 1987;27:980-7.
23 Bachulis Bl, Long WB, Hynes GD, Johnson MC. Clinical indications for cervical spine radiographs in the traumatized patient. *Am J Surg* 1987;153:473-7.

(Accepted 14 March 1997)

Alcohol consumption and cognitive performance in a random sample of Australian soldiers who served in the second world war

O F Dent, M R Sulway, G A Broe, H Creasey, S C Kos, A F Jorm, C Tennant, M J Fairley

Abstract

Objective: To examine the association between the average daily alcohol intake of older men in 1982 and cognitive performance and brain atrophy nine years later.
Subjects: Random sample of 209 Australian men living in the community who were veterans of the second world war. Their mean age in 1982 was 64.3 years.
Main outcome measures: 18 standard neuropsychological tests measuring a range of intellectual functions. Cortical, sylvian, and vermian atrophy on computed tomography.
Results: Compared with Australian men of the same age in previous studies these men had sustained a high rate of alcohol consumption into old age. However, there was no significant correlation, linear or non-linear, between alcohol consumption in 1982 and results in any of the neuropsychological tests in 1991; neither was alcohol consumption associated with brain atrophy on computed tomography.
Conclusion: No evidence was found that apparently persistent lifelong consumption of alcohol was related to the cognitive functioning of these men in old age.

Introduction

Epidemiological evidence of the differential age related risks and possible benefits of drinking has recently prompted reassessment of sensible levels of drinking and their appropriateness for different age groups.[1 2] The potential effects of alcohol consumption on cognitive performance in elderly people are of particular interest in this regard.[3]

The Australian National Health and Medical Research Council set its recommendations for responsible alcohol consumption on the basis of a wide range of consequences of excessive drinking among younger adults rather than elderly people.[4] However, the national health survey has shown that 8.5% of Australian men aged 65-74 years who consumed alcohol drank at the hazardous level (40-60 g per day) and 5.7% at the harmful level (>60 g per day).[5] We examined the association between the average daily alcohol consumption in 1982 of a random sample of Australian male veterans of the second world war and performance in a range of intellectual tests and the degree of brain atrophy on computed tomography nine years later.

Department of Sociology, Australian National University, Canberra, ACT 0200, Australia
O F Dent, *reader*
Centre for Education and Research on Ageing, Concord Hospital, Sydney, NSW 2319, Australia
M R Sulway, *research officer*
G A Broe, *professor*
H Creasey, *senior specialist*

BMJ 1997;314:1655–7

continued over

Departments of Radiology and Psychiatry, Concord Hospital, Sydney, NSW 2139
S C Kos, *senior specialist in radiology*
M J Fairley, *staff psychiatrist*
National Health and Medical Research Council Social Psychiatry Research Unit, Australian National University, Canberra, ACT 0200
A F Jorm, *deputy director*
University of Sydney Academic Psychiatric Unit, Royal North Shore Hospital, Sydney, NSW 2065
C Tennant, *professor*
Correspondence to: Dr Dent.

Table 1 Correlations between alcohol consumption in 1982 and neuropsychological test scores in 1991

	Bivariate analysis				Quadratic analysis		
Test	**No of men**	**ρ**	**P value**	***r* (95% CI)**	**Multiple *r***	**F**	**P value**
Wechsler adult intelligence scale (revised 1981)[8]:	209	–0.06	0.41	–0.03 (–0.16 to 0.11)	0.12	1.5	0.22
Digit span	209	0.07	0.34	0.10 (–0.03 to 0.24)	0.13	1.7	0.18
Block design	207	0.04	0.56	0.04 (–0.10 to 0.17)	0.05	0.3	0.78
Similarities	208	–0.08	0.25	–0.08 (–0.21 to 0.06)	0.11	1.3	0.29
Wechsler memory scale (revised 1987)[9]:							
Immediate	209	0.04	0.57	0.04 (–0.10 to 0.18)	0.05	0.3	0.74
Delayed	208	0.08	0.26	0.08 (–0.05 to 0.22)	0.08	1.4	0.23
Benton visual retention test (1974)[10]:							
Correct	206	0.00	0.97	0.00 (–0.14 to 0.14)	0.04	0.2	0.83
Errors	205	0.01	0.88	0.03 (–0.11 to 0.17)	0.04	0.2	0.84
Raven's coloured progressive matrices (1984)[12]	207	–0.04	0.53	–0.05 (–0.18 to 0.09)	0.05	0.2	0.79
Rey auditory verbal learning test (1964)[13]:							
Trial 5	207	–0.03	0.65	–0.02 (–0.15 to 0.12)	0.03	0.1	0.88
List B recall	207	–0.10	0.14	–0.08 (–0.22 to 0.06)	0.11	1.3	0.28
List A recall	207	0.04	0.61	0.07 (–0.07 to 0.20)	0.08	0.7	0.51
Controlled oral word association test (1983)[14]	207	–0.01	0.85	–0.02 (–0.15 to 0.12)	0.08	0.7	0.49
Boston naming test (1983)[15]	205	–0.08	0.25	–0.05 (–0.19 to 0.09)	0.05	0.3	0.74
National adult reading test (1982)[16]	200	–.04	0.59	0.04 (–0.10 to 0.18)	0.12	1.4	0.24
Reaction time:							
Simple	202	–0.05	0.45	–0.02 (–0.15 to 0.12)	0.09	0.8	0.43
Delayed	202	–0.02	0.75	–0.02 (–0.16 to 0.12)	0.13	1.7	0.19
Subscale of Halstead Reitan battery (1958)[18]	182	–0.10	0.17	–0.09 (–0.24 to 0.05)	0.11	1.1	0.32

Subjects and methods

This study arose from a project on morbidity among war veterans which began in 1982 but did not include assessment of cognitive performance at that time.[6] Participants were chosen randomly from former members of the Australian army who had fought in the second world war and were living in Sydney.[6] The response rate was 87%. Survivors were traced in 1991 for the second phase of the study.

Self reported average daily alcohol intake was assessed in both 1982 and 1991 by the quantity-frequency method.[7] In 1991, 18 standard neuropsychological tests were used to assess several aspects of cognitive performance (see table 1).[8-18]

Non-contrast computed tomography was performed on 201 veterans in 1991 using a Siemens Somatom DR3 scanner (Germany) at a slice thickness of 8 mm. Cortical, sylvian, and vermian atrophy were recorded as none, slight, moderate, or severe by a radiologist (SCK) who was unaware of subjects' alcohol consumption or neuropsychological test scores.

As the distribution of alcohol consumption was strongly positively skewed, correlations with the neuropsychological tests were assessed by both the Spearman rank coefficient (ρ) on the original values of alcohol intake and the Pearson product moment coefficient (*r*) applied to log transformed values. The possibility of a non-linear association between alcohol consumption and results of the cognitive function tests was examined by fitting quadratic curves. Contingency tables and the χ^2 test were used to examine the significance of differences between percentages. The *t* test for related samples was used to assess the significance of change in average daily consumption between 1982 and 1991.

To estimate the power of the sample we regarded an *r* of at least 0.2 as the minimum effect size we would wish to identify. At this effect size and with significance set at 0.05 the power of the sample of 209 would be at least 0.8.[19]

Table 2 Moderate or severe cortical, sylvian, and vermian atrophy in 1991 by alcohol consumption in 1982. Values are numbers (percentages; 95% confidence intervals) of subjects

		Level of drinking				
Atrophy	**Non-drinker (n=31)**	**Safe (n=101)**	**Hazardous (n=32)**	**Harmful (n=37)**	**χ^2**	**P value**
Cortical	23 (74; 59 to 90)	72 (71; 63 to 80)	25 (78; 64 to 92)	28 (76; 62 to 90)	0.7	0.9
Sylvian	26 (84; 71 to 97)	81 (80; 72 to 88)	26 (81; 68 to 95)	31 (84; 72 to 96)	0.4	0.9
Vermian	14 (45; 28 to 63)	56 (55; 46 to 65)	17 (53; 36 to 70)	23 (62: 47 to 79)	2.0	0.6

Results

Of the 342 veterans studied in 1982, 96 had died by 1991, 10 had moved away from Sydney, eight could not be located or were too ill to attend, and 19 refused, leaving 209 consenting participants. In 1982 the average daily consumption of the 178 veterans (85%) who drank alcohol at least once a week ranged from 2 g to 129 g ethanol (mean 38.2 g); 60% (106/178) drank at the safe level, 19% (34/178) at the hazardous level, and 21% (38/178) at the harmful level, the proportions drinking at the hazardous and harmful levels being appreciably higher than those in the Australian male population aged 55 to 74 (9.8% and 7.7% respectively).[5] The average period over which they had drunk alcohol ranged from 11 to 56 years (mean 44 (SD 5.6) years).

No significant correlation was found between average daily alcohol consumption in 1982 and any of the cognitive performance measures in 1991 (table 1), and in no case was there evidence of a significant non-linear association. There was no correlation between age and alcohol consumption ($\rho = -0.05$, $P = 0.39$).

The proportion of drinkers fell from 85% in 1982 to 67% in 1991 (McNemar $\chi^2 = 29.76$, $P < 0.0001$). Among those who drank in 1982 and were still drinking in 1991, average daily consumption fell from 43.7 g to 30.5 g (paired $t = 6.04$, $P < 0.001$). However, among

drinkers the correlation between consumption in 1982 and 1991 was moderately strong ($r = 0.58$, $P < 0.0001$), and in 1991, 65% (91/140) drank at the safe level, 26% (37/140) at the hazardous level, and 9% (12/140) at the harmful level. There was no association, linear or non-linear, between alcohol consumption in 1991 and any of the measures of cognitive performance.

The proportions of veterans with moderate or severe cortical, sylvian, or vermian atrophy on computed tomography in 1991 did not differ significantly between those who were non-drinkers and drinkers in 1982; neither was there any difference in relation to alcohol consumption (table 2).

Our failure to find an association between alcohol consumption and cognitive performance or brain atrophy in these men could be a result of differential mortality or loss to follow up. However, the proportion of men who died or were not included in the 1991 survey did not differ significantly according to alcohol consumption in 1982, whether or not stratified by age.

Discussion

Several recent studies of elderly social drinkers have found no association or only weak associations between cognitive performance and alcohol consumption measured either concurrently or retrospectively.[20-25] Furthermore, any correlations found have tended to disappear when other factors were controlled statistically. We know of only one other prospective study of the effects of alcohol consumption on cognitive performance in elderly people living in the community.[26]

The elderly veterans we studied were originally recruited in 1982 for another purpose that did not entail measuring cognitive performance, so we could not use cognitive decline between the two studies as an outcome variable. Instead we considered whether apparently high lifelong alcohol consumption during adulthood among a comparatively large number of men in our sample was predictive of diminished cognitive performance in old age. It is commonly believed that many Australian veterans of the second world war have continued to drink comparatively heavily during the 50 years since the war. This continued heavy drinking might be the result of habits acquired during the war, mateship, a desire to suppress memories of wartime experiences, or the atmosphere of the ubiquitous clubs of the Returned Serviceman's League. Compared with Australian men of the same age, many of the veterans in our study had sustained a high rate of alcohol consumption into old age. If alcohol consumption were clearly associated with cognitive performance then the relation would be expected in this group of men. However, we found no association between consumption in 1982 and brain atrophy on computed tomography or several aspects of cognitive functioning almost a decade later in a group whose cognitive performance was within the normal range for people of their age. There is no basis here for predicting future cognitive performance from level of alcohol consumption in older men, even though many had apparently experienced a lifetime of persistent, moderately heavy consumption.

Key messages

- The effects of lifelong alcohol consumption on cognitive function are of interest in view of recent changes to recommendations for sensible drinking
- In this study of Australian veterans of the second world war, alcohol consumption measured in 1982 was not associated with performance nine years later in 18 standard neuropsychological tests measuring a range of intellectual functions
- Computed tomography showed no difference in brain atrophy between drinkers and non-drinkers and no difference according to alcohol consumption in 1982
- Apparently persistent lifelong consumption of alcohol and the level of intake seemed not to have any impact on cognitive performance among these men in old age

We thank the veterans who gave their time willingly to participate.

Funding: The first phase of this project was funded by the Australian Department of Veterans' Affairs and the second by the Australian National Health and Medical Research Council.

Conflict of interest: None.

1 Jackson R, Beaglehole R. Alcohol consumption guidelines: relative safety vs absolute risks and benefits. *Lancet* 1995;346:716.
2 Edwards G. Sensible drinking. *BMJ* 1996;312:1.
3 Dufor MC, Archer L, Gordis E. Alcohol and the elderly. *Clin Geriatr Med* 1992;8:127-41.
4 Pols RG, Hawks DV. *Is there a safe level of daily alcohol consumption for men and women?* Canberra: Australian Government Publishing Service, 1992.
5 Castles I. *1989-90 National health survey: summary of results.* Canberra: Australian Bureau of Statistics, 1991. (No 4364.0.)
6 Goulston KJ, Dent OF, Chapuis PH, Chapman YE, Smith CI, Tait AD, *et al.* Gastrointestinal morbidity among world war II prisoners of war: 40 years on. *Med J Aust* 1985;143:6-10.
7 Castles I. *1989-90 National health survey users' guide.* Canberra: Australian Bureau of Statistics, 1991. (No 4363.0.)
8 Wechsler D. *Wechsler adult intelligence scale—revised.* New York: Psychological Corporation, 1981.
9 Wechsler D. *WMS-R: Wechsler memory scale—revised manual.* New York: Psychological Corporation, 1987.
10 Benton AL. *Revised visual retention test.* New York: Psychological Corporation, 1974.
11 Raven JC. *Guide to using the coloured progressive matrices sets A, AB, B.* London: HK Lewis, 1965.
12 Raven JC, Court JH, Raven J. *Manual for Raven's progressive matrices and vocabulary scales. Coloured progressive matrices.* London: HK Lewis, 1984.
13 Rey A. *L'examen clinique en psychologie.* Paris: Presses Universitaires de France, 1964.
14 Lezak MD. *Neuropsychological assessment.* 2nd ed. New York: Oxford University Press, 1983.
15 Kaplan E, Goodglass H, Weintraub S. *Boston naming test.* Philadelphia: Lea and Febiger, 1983.
16 Nelson HE. *The national adult reading test (NART) manual.* Windsor: NFER-Nelson, 1982.
17 Reid WGJ, Broe GA, Hely MA, Morris JGL, Williamson PM, O'Sullivan DJ, *et al.* The neuropsychology of de novo patients with idiopathic Parkinson's disease: the effects of age of onset. *Int J Neuroscience* 1989;48:205-17.
18 Reitan RM. Validity of the trail making test as an indicator of organic brain damage. *Percept Motor Skills* 1958;8:271-6.
19 Cohen J. *Statistical power analysis for the behavioral sciences.* New York: Academic Press, 1977.
20 Bowden SC. Brain impairment in social drinkers? No cause for concern. *Alcohol Clin Exp Res* 1987;11:407-10.
21 Goodwin JS, Sanchez CJ, Thomas P, Hunt C, Garry PJ, Goodwin JM. Alcohol intake in a healthy elderly population. *Am J Public Health* 1987;77:173-7.
22 Page RD, Cleveland MF. Cognitive dysfunction and aging among male alcoholics and social drinkers. *Alcohol Clin Exp Res* 1987;11:376-84.
23 Scherr PA, Albert MS, Funkenstein HH, Cook NR, Hennekens CH, Branch LG, *et al.* Correlates of cognitive function in an elderly community population. *Am J Epidemiol* 1988;128:1084-101.
24 Waugh M, Jackson M, Fox GA, Hawke SH, Tuck RR. Effects of social drinking on neuropsychological performance. *Br J Addict* 1989;84:659-67.
25 Iliffe S, Haines A, Booroff A, Goldenberg E, Morgan P, Gallivan S. Alcohol consumption by elderly people: a general practice survey. *Age Ageing* 1991;20:120-3.
26 Herbert LE, Scherr PA, Beckett LA, Albert MA, Rosner B, Taylor JO, *et al.* Relation of smoking and low-to-moderate alcohol consumption to change in cognitive function: a longitudinal study in a defined community of older persons. *Am J Epidemiol* 1993;137:881-91.

(Accepted 8 April 1997)

Short term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project

K Katsouyanni, G Touloumi, C Spix, J Schwartz, F Balducci, S Medina, G Rossi, B Wojtyniak, J Sunyer, L Bacharova, J P Schouten, A Ponka, H R Anderson

Department of Hygiene and Epidemiology, University of Athens Medical School, Athens 115 27, Greece
K Katsouyanni, *associate professor*
G Touloumi, *research fellow*
GSF-Forschungszentrum für Umwelt und Gesundheit, Institut für Epidemiologie Postfach 1129 Neuherberg, Germany
C Spix, *statistician*
Harvard School of Public Health, Environmental Epidemiology Program, Boston, MA 02115, USA
J Schwartz, *associate professor*
Faculté de Medicine, Université de Grenoble, Department of Public Health, Domaine de la Merci, F-38706, La Tronche, Cedex, France
F Balducci, *research fellow*
Observatoire Regional de la Santé, Paris, ORS Ile de France, 75732 Paris Cedex 15, France
S Medina, *senior researcher*
Institute of Clinical Physiology, National Research Council, IFC-CNR, 56100 Pisa, Italy
G Rossi, *researcher*
National Institute of Hygiene, 00-791 Warsaw, Poland
B Wojtyniak, *senior researcher*

BMJ 1997;314:1658–63

continued over

Abstract

Objectives: To carry out a prospective combined quantitative analysis of the associations between all cause mortality and ambient particulate matter and sulphur dioxide.
Design: Analysis of time series data on daily number of deaths from all causes and concentrations of sulphur dioxide and particulate matter (measured as black smoke or particles smaller than 10 μm in diameter (PM_{10})) and potential confounders.
Setting: 12 European cities in the APHEA project (Air Pollution and Health: a European Approach).
Main outcome measure: Relative risk of death.
Results: In western European cities it was found that an increase of 50 μg/m^3 in sulphur dioxide or black smoke was associated with a 3% (95% confidence interval 2% to 4%) increase in daily mortality and the corresponding figure for PM_{10} was 2% (1% to 3%). In central eastern European cities the increase in mortality associated with a 50 μg/m^3 change in sulphur dioxide was 0.8% (−0.1% to 2.4%) and in black smoke 0.6% (0.1% to 1.1%). Cumulative effects of prolonged (two to four days) exposure to air pollutants resulted in estimates comparable with the one day effects. The effects of both pollutants were stronger during the summer and were mutually independent.
Conclusions: The internal consistency of the results in western European cities with wide differences in climate and environmental conditions suggest that these associations may be causal. The long term health impact of these effects is uncertain, but today's relatively low levels of sulphur dioxide and particles still have detectable short term effects on health and further reductions in air pollution are advisable.

Introduction

The adverse health effects of episodes of severe air pollution were established over 40 years ago by investigations in Europe and North America.[1] Since then the levels of particles and sulphur dioxide, which characterised these episodes, have fallen below the limits considered safe for human health in the 1980s.[2]

More recent studies, mainly from the United States,[3 4] have found associations between mortality and other health indicators and levels of outdoor suspended particulate matter that are well within existing air quality guidelines and standards. These studies have had a considerable impact on scientific dialogue and standards.[5 6] However, little information exists about the possible effects of the current concentrations of sulphur dioxide.

Apart from data from a few European studies which used various methodological approaches,[6-10] little is known about the short term effects of air pollution on mortality in Europe. There are many differences between Europe and the United States, and within Europe itself, which might influence the health effects of air pollution; these include emission sources, pollution mixes, climate, lifestyle, and the underlying health of the population.

A multicity analysis of the short term health effects of air pollution on mortality and hospital emergency admissions was initiated within the European Union Environment 1991-94 Programme (Air Pollution and Health: a European Approach: the APHEA project). The study investigated the effects of several air pollutants in 15 European cities in 10 countries.[11] This paper reports the combined results from the 12 cities which had data available for investigating the effects of sulphur dioxide and particulate matter on daily mortality.

Methods

Table 1 shows the cities studied, the size of populations, the concentrations of the available pollutants, the length of study, and the mean daily number of deaths. Particulate matter was measured either as black smoke (determining the intensity of the black stain produced mainly by particles below 4 μm in diameter) or as particles below 10 μm in diameter (PM_{10}). Under the protocol the data for each city were analysed separately but with a standardised approach to data eligibility and statistical analysis.[12 13] About half of the data from individual cities used in this paper have been published previously.[14 15]

The daily data from the cities were analysed by time series methods.[13] A specific standardised method was applied to control for confounding while at the same time allowing the flexibility to take account of local characteristics. This procedure included modelling all potential confounders (seasonal and long term patterns, daily temperature, humidity, day of the week, holidays, influenza epidemics, and other unusual events), choosing the "best" air pollution models, and applying diagnostic tools to check the adequacy of the models. For each pollutant the best one day measurement chosen from lags 0 (same day) to 3 and the best average indicating cumulative exposure (up to four consecutive days) were selected by each centre.

Several previously decided pollutant transformations were tested. Generally, in cleaner cities linear terms for the pollutants fitted the data best. When log transformations had the best fit, additional models were fitted with linear pollutant terms, restricting the analysis to days when the pollutant did not exceed 200 μg/m^3. The final analysis was done with autoregressive Poisson models, allowing for overdispersion

Table 1 Duration of analysis, population, and pollutant data for cities contributing to analysis of total daily mortality

Town	Length of study	Population* (x 1000)	No of deaths/day (SD)	Black smoke ($\mu g/m^3$) percentiles		Sulphur dioxide ($\mu g/m^3$) percentiles		Particulate matter† ($\mu g/m^3$) percentiles	
				50	90	50	90	50	90
Athens	1987-91	2000	35 (6)	73	146	45	86	—	—
Barcelona	1986-92	1700	46 (9)	40	76	41	77	85	116
Bratislava	1987-91	443	11 (3)	—	—	13	50	39	95
Cracow	1977-89	740	18 (5)‡	73	247	74	170	—	—
Cologne	1975-85	977	30 (6)	—	—	44	96	34	69
Lodz	1977-90	848	28 (6)‡	57	151	46	123	—	—
London	1987-91	7200	199 (21)‡	13	23	29	45	—	—
Lyons	1985-90	410	8 (3)‡	—	—	37	86	33	64
Milan	1980-89	1500	32 (7)‡	—	—	66	293	66	137
Paris	1987-92	6140	128 (15)‡	26	56	23	59	47	81
Poznan	1983-90	575	18 (4)‡	34	92	41	131	—	—
Wroclaw	1979-89	637	14 (4)‡	54	141	29	83	—	—

*Numbers refer to the populations covered by the data collection.†Particles <13 μm for Paris and Lyons and <7 μm for Cologne; total suspended particles (TSP) for the other cities, converted to particles <10 μm (PM_{10}) with the formula: PM_{10}=TSP*0.55.‡Excluding deaths from external causes.

and autocorrelation where necessary. Modification of the effect by season and the levels of other pollutants was also tested by appropriate models. Effects were reported by each city as partial regression coefficients, standard errors, and covariances (where needed) from the final Poisson models. Details of the analytical methods used for each city's data have been published.[12 13]

The city specific estimates of effect for each model were combined quantitatively. The summary estimates were weighted means of the regression coefficients, with weights inversely proportional to local variances (fixed effects model).[16] All available coefficients were included in the analysis. Homogeneity of the coefficients was tested with a χ^2 test under the fixed effects hypothesis. If significant heterogeneity was present ($P<0.05$), its determinants were investigated by using a predefined list of explanatory variables which included the levels of the pollutant evaluated and the levels of other pollutants (annual mean or seasonal mean and correlations between pollutants); meteorological factors (annual mean or seasonal temperature and humidity); accuracy of measurements of air pollutants (number of monitoring sites, correlations between measurements from different sites); health of the population (age standardised mortality, proportion of elderly people); smoking prevalence; geographical differences (north-south, east-west). A random effects model was also applied in case of significant unexplained heterogeneity. Under this model the between cities variance is added to the estimates of the local variances as a way of quantifying the inherent greater uncertainty of heterogeneous results.

Results

The participating cities showed substantial variation in air pollution mixtures and concentrations and in geographical distribution, seasonal patterns, and meteorological and climatic conditions.[11] Mean winter concentrations of sulphur dioxide varied from 30 to 330 $\mu g/m^3$ and of black smoke from 10 to 290 $\mu g/m^3$; the mean summer concentrations of nitrogen dioxide varied from 60 to 205 $\mu g/m^3$ and of ozone from 55 to 165 $\mu g/m^3$. Typical patterns of winter and summer type smog were observed, with some cities having particularly high winter type smog, dominated either by sulphur dioxide (for example, Milan) or particles (for example, Cracow) while others had relatively high air pollution of both types (for example, Athens). The average temperature and humidity in the winter ranged from -2^oC to 10^oC and from 67% to 89% respectively and in the summer from 17^oC to 26^oC and from 50% to 76%.[11]

Table 2 shows the pooled estimates of relative risk associated with a 50 $\mu g/m^3$ change in 24 hour pollutant levels and the tests for heterogeneity for black smoke, PM_{10}, and sulphur dioxide for one day and cumulative effects. Significant heterogeneity was found for the effects of sulphur dioxide and black smoke. The attempt to explain this heterogeneity by the variables described in the Methods section showed that only the separation between western and central eastern European cities resulted uniformly in more homogeneous subgroups. However, significant heterogeneity still remained for the effect of sulphur dioxide in western cities. Overall, an increase of 50 $\mu g/m^3$ in the one day pollutant levels was associated with an increase in the daily mortality of 3% for sulphur dioxide, 3% for

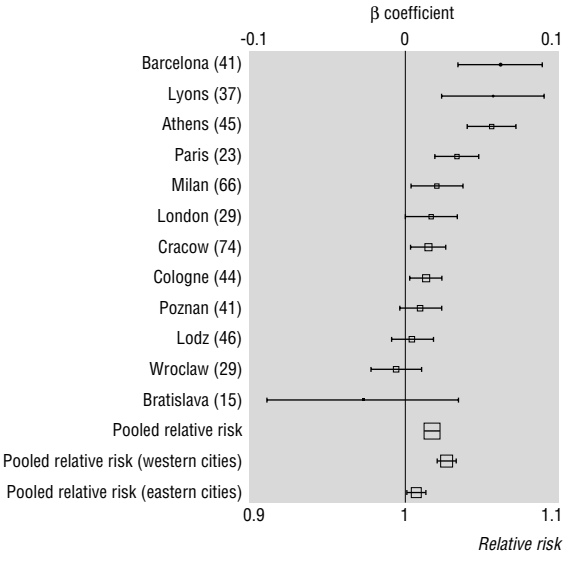

Fig 1 Estimated individual city and pooled relative risks of mortality associated with increase of 50 $\mu g/m^3$ in sulphur dioxide concentration. Numbers in parentheses are median value of pollutant, and the size of the point representing each relative risk is inversely proportional to its variance

Institut Municipal D' Investigacio Medica, E-08003, Barcelona, Spain
J Sunyer, *scientist*
National Centre for Health Promotion, 82007 Bratislava, Slovakia
L Bacharova, *internist*
Department of Epidemiology and Statistics, University of Groningen, Groningen 9713, Netherlands
J P Schouten, *associate professor*
Helsinki City Centre of the Environment, Environmental Health Unit, 00530 Helsinki, Finland
A Ponka, *head*
Department of Public Health Sciences, St George's Hospital Medical School, London SW17 0RE
H R Anderson, *professor*
Correspondence to: Dr Katsouyanni.

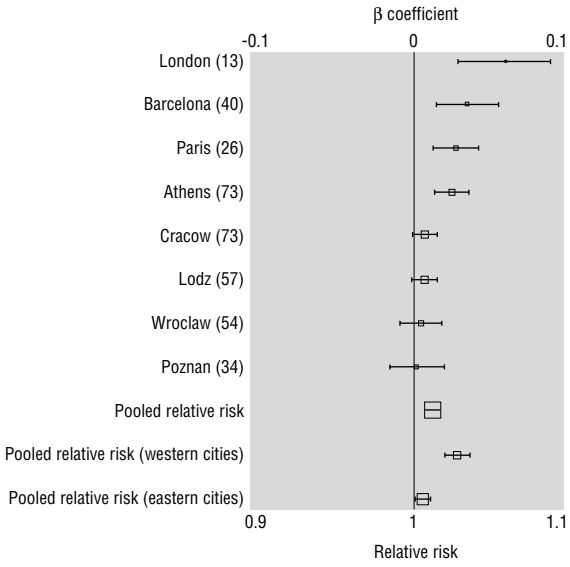

Fig 2 Estimated individual city and pooled relative risks of mortality associated with increase of 50 μg/m^3 in black smoke levels. Numbers in parentheses are median value of pollutant, and the size of the point representing each relative risk is inversely proportional to its variance

black smoke, and 2% for PM_{10} in the western European cities. The corresponding figures for central eastern cities were 1% for sulphur dioxide and black smoke (4% for PM_{10} in Bratislava, the only city with results available). The cumulative effects were consistent with the one day estimates but the relative risks were larger, as would be expected since they reflect the same day and the lagged effects of air pollution.

Figures 1-3 show the relative risks for 50 μg/m^3 change in the pollutant levels for each city as well as the pooled estimates (also separately for central eastern and western European cities). When two pollutant models were fitted, including both sulphur dioxide and black smoke or PM_{10}, the regression coefficients for the effects of sulphur dioxide and black smoke both decreased by 32%. The relative risks (95% confidence intervals) became 1.023 (1.007 to 1.039) and 1.020 (1.000 to 1.040) for an increase of 50 μg/m^3 for sulphur dioxide and black smoke respectively.

Table 3 shows the pooled estimates separately for summer and winter for western and central eastern European cities. The effects were stronger in the summer per unit increase in the pollutant in the western cities. The difference between seasons was significant for PM_{10}. All the pollutants analysed had highest concentrations during winter.

Table 4 shows the pooled estimated effects for sulphur dioxide for days with high or low particle levels and the corresponding black smoke effects for days with high or low sulphur dioxide concentrations in western European cities. The effects of sulphur dioxide and black smoke were similar for days with low or high levels of the other pollutant and the same as their overall effect.

Discussion

Air Pollution and Health: a European Approach is a systematic attempt to analyse time series data on air pollution and health in a standardised way on a large scale. The cities included in the project covered over 23 million people and represented various environmental and climatic conditions, offering an important opportunity for assessing the consistency of the association between air pollution and mortality. The estimated effects are for moderate and low levels of pollution (particulate matter and sulphur dioxide below 200 μg/m^3), which are the most relevant exposures in current situations.

Table 2 Estimated pooled relative risks and 95% confidence intervals for 50 μg/m^3 change in 24 hour pollutant levels* from Poisson autoregressive models of sulphur dioxide, particulate matter (PM_{10}), and black smoke on total mortality

	All cities			Western cities			Central eastern cities		
Pollutant	**No of cities**	**Relative risk (95% CI)**	**P value†**	**No of cities**	**Relative risk (95%CI)**	**P value†**	**No of cities**	**Relative risk (95% CI)**	**P value†**
Sulphur dioxide									
1 Day:									
Fixed effects model	12	1.020 (1.015 to 1.024)	<0.0001	7	1.029 (1.023 to 1.035)	<0.001	5	1.008 (0.993 to 1.024)	0.25
Random effects model‡	—	—	—	7	1.035 (1.020 to 1.050)	—	—	—	—
Cumulative§:									
Fixed effects model	12	1.023 (1.017 to 1.028)	<0.0001	7	1.032 (1.024 to 1.040)	<0.0001	5	1.011 (1.002 to 1.019)	0.04
Random effects model‡	—	—	—	7	1.040 (1.018 to 1.062)	—	5	1.008 (0.993 to 1.023)	—
Black smoke									
1 Day	8	1.013 (1.009 to 1.017)	0.08	4	1.029 (1.021 to 1.037)	0.34	4	1.006 (1.001 to 1.011)	0.99
Cumulative§	8	1.014 (1.009 to 1.020)	<0.001	4	1.031 (1.022 to 1.040)	0.08	4	1.004 (0.997 to 1.011)	0.40
PM_{10}¶									
1 Day	6	1.022 (1.013 to 1.031)	0.53	5	1.021 (1.012 to 1.030)	0.58	1	1.043 (1.003 to 1.085)	—
Cumulative§	6	1.021 (1.012 to 1.031)	0.43	5	1.031 (1.022 to 1.040)	0.08	1	1.049 (0.996 to 1.105)	—

*Averaging time.
†P value from χ^2 for heterogeneity.
‡Used only when there was significant heterogeneity.
§Average of 2 to 4 consecutive days, including the day of recorded mortality
¶PM_{10}: PM_{13} for Paris and Lyon and PM_7 for Cologne; total suspended particles (TSP) for other cities were converted to PM_{10} by the formula PM_{10} = TSP*0.55.

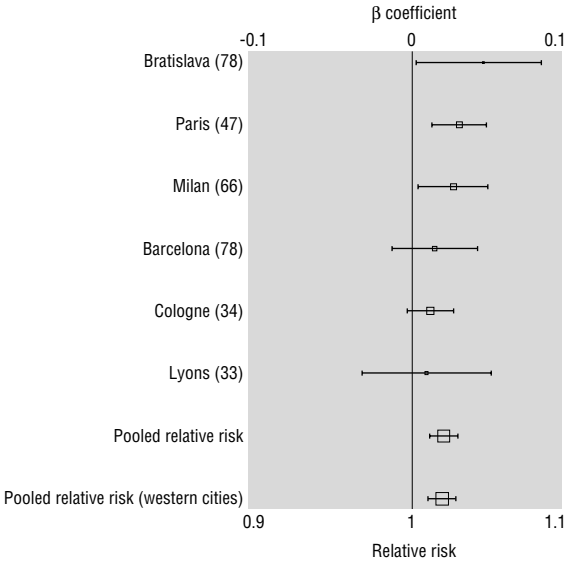

Fig 3 Estimated individual city and pooled relative risks of mortality associated with an increase of 50 μg/m^3 in particulate matter (PM_{10}) levels. Numbers in parentheses are median value of pollutant, and the size of the point representing each relative risk is inversely proportional to its variance

The project planned a combined quantitative summary of results from all the cities. In this respect it differs from most meta-analyses, which are started only after a substantial number of papers on a specific hypothesis have been published. It follows that several common problems affecting meta-analyses do not apply here: there was no publication bias (all the results were included), no selection bias (in the sense that all cities which could provide data and enter the project were included long before the results were known), and all the necessary coefficients were available.

We found a clear significant effect of moderate to low levels of sulphur dioxide and particles (< 200 μg/m^3) in western European cities which seems independent for the two pollutants. In the United States meta-analyses focusing on the effects of particles (using total suspended particles and PM_{10} as indicators) estimated that an increase of 100 μg/m^3 total suspended particles (about 50 μg/m^3 PM_{10}) is associated with a relative risk of daily deaths of 1.05 to 1.06.[3 4] The relative risks for PM_{10} in the individual studies (for about 50 μg/m^3 change) ranged from 1.03 to 1.08; the largest estimates were in the areas with the lowest sulphur dioxide concentrations (Utah Valley; Kingston, Tennessee; St Louis, Missouri). In Philadelphia, where there are moderate sulphur dioxide concentrations, detailed recent analyses indicate that both total suspended particles and sulphur dioxide affect death rates.[17] These results have been shown to be insensitive to different methods of analysis.[18]

In Europe, two earlier studies indicated that the effect of sulphur dioxide is more pronounced than that of particles (measured as black smoke), but these results cannot be readily interpreted as relative risks.[6 7] Analyses of daily mortality and air pollution in the Netherlands showed no effect of sulphur dioxide (which was at very low concentrations) on mortality but reported a significant association with black smoke.[19] Data from Athens for the five years before this study indicated independent effects of both sulphur dioxide and black smoke on total mortality.[20]

The size of the effect of particles (either PM_{10} or black smoke) that we found is compatible with those reported in the United States and in other studies, although at the lower end of the range of relative risks. We found, however, that the effect of sulphur dioxide on mortality, which has not been so thoroughly studied elsewhere, was consistent with and of a similar size to the effect of particles. Furthermore, inclusion of estimates from recent analyses in five additional western European cities[19 21 22] results in a pooled relative risk associated with a 50 μg/m^3 change in the pollutant level of 1.028, 1.025, and 1.030 for sulphur dioxide, PM_{10}, and black smoke respectively. These values are similar to our original estimates.

The reason for our results for particles being on the lower side of the range reported in the United States may be the more complex mixtures of pollutants that we observed. Thus, in a place where there are high levels of particles in the absence of sulphur dioxide the particle effect may be higher because there is no competing pollutant to deplete the pool of people at high risk. Alternatively, the distribution of the size and toxicity of the particles may differ or the proportion of susceptible individuals may be larger in the American cities studied.

Table 3 Estimated pooled relative risks and 95% confidence intervals for 50 μg/m^3 change in 24 hour pollutant levels from Poisson autoregressive models of sulphur dioxide, particulate matter (PM_{10}), and black smoke on total mortality by season

	Warm season		Cold season	
Pollutant	**Relative risk (95% CI)**	**P value***	**Relative risk (95% CI)**	**P value***
Sulphur dioxide				
Western:				
Fixed effects model	1.040 (1.028 to 1.053)	0.35	1.024 (1.071 to 1.031)	<0.001
Random effects model†	—	—	1.033 (1.015 to 1.051)	—
Central eastern:				
Fixed effects model	1.011 (0.999 to 1.023)	0.003	1.008 (1.002 to 1.015)	0.05
Random effects model†	1.008 (0.979 to 1.037)		1.009 (0.997 to 1.020)	—
PM_{10}‡				
Western	1.043 (1.027 to 1.060)	0.31	1.010 (0.998 to 1.022)	0.81
Central eastern	1.053 (0.994 to 1.116)	—	1.019 (0.941 to 1.107)	—
Black smoke				
Western	1.040 (1.024 to 1.057)	0.27	1.024 (1.016 to 1.032)	0.50
Central eastern	1.013 (1.002 to 1.024)	0.40	1.005 (0.998 to 1.011)	0.96

*P value from χ^2 for heterogeneity.
†Given only when there was significant heterogeneity.
‡PM_{10}: PM_{13} for Paris and Lyon and PM_7 for Cologne; total suspended particles (TSP) for the other cities were converted to PM_{10} by the formula: PM_{10} = TSP*0.55.

Table 4 Estimated pooled relative risks of deaths from all causes (95% confidence intervals) for 50 μg/m^3 change in 24 hour pollutant levels from Poisson autoregressive models of sulphur dioxide and black smoke by high and low levels of other pollutant in western European cities

	Days with lower level of other pollutant		Days with high level of other pollutant*	
Pollutant	**Relative risk (95% CI)**	**P value†**	**Relative risk (95% CI)**	**P value†**
Sulphur dioxide				
Fixed effects model	1.029 (1.016 to 1.043)	0.11	1.030 (1.022 to 1.038)	0.002
Random effects model‡	—	—	1.034 (1.019 to 1.050)	—
Black smoke				
Fixed effects model	1.028 (1.013 to 1.043)	0.10	1.027 (1.019 to 1.035)	0.15

*Above the city's median level.†P value from χ^2 for heterogeneity. ‡Random effects model, only when there is significant heterogeneity.

Independent effects

There is evidence from three different approaches that the effects of the two pollutants investigated are independent. Firstly, if the effect of one pollutant was a surrogate of the other, then ranking the effects estimated for one pollutant by the mean level of the other—for example, plotting the sulphur dioxide coefficient for all cities ranked by their black smoke level—should show a monotonic trend in the size of the effect. This was not found for either sulphur dioxide or any particle measurement. Secondly, models for the two pollutants during days with low and high levels of the other pollutant were fitted, and the pooled results indicated independent effects. Thirdly, the results of the two pollutant models showed a similar moderate decrease in the effects of both pollutants, which remained significant.

It is possible, however, that either or both these pollutants may be surrogates for other unmeasured substances. In fact, the exposure variables that we used are proxies for personal exposure correlated with day to day differences in personal exposure to particles and sulphur dioxide or, generally, to primary combustion related pollution. If the responsible pollutants must be those to which individuals are exposed for longer periods during a day, then the pollutants which penetrate and remain suspended indoors are strong candidates. Fine particles (with a diameter <2.5 μm), and in particular sulphates, have been found suspended indoors in the absence of indoor sources.[23] The stronger correlation of mortality with black smoke rather than PM_{10} found in this study may reflect the fact that black smoke consists only of fine particles or may reflect a relatively greater toxicity of diesel related pollution, which is the major source of black particles in many cities. The effects of both pollutants studied are not likely to be confounded by ozone or nitrogen dioxide concentrations as their correlation with ozone is low and the observed effects of nitrogen dioxide on mortality were relatively weak and largely confounded by particle levels.[24]

How health is affected

Our results are consistent with the effects of particles and sulphur dioxide on cardiovascular and respiratory mortality found within the APHEA project. Specifically, in five western European cities an increase in sulphur dioxide concentrations of 50 μg/m^3 was associated with a 4% and 5% rise in cardiovascular and respiratory mortality respectively; the corresponding figures for black smoke were 2% and 4% (unpublished data).

The biological mechanism by which exposure to particulate matter may increase mortality is not well understood but has received considerable attention.[25] Acidic ultrafine particles may provoke alveolar inflammation causing acute changes in blood coagulability and release of mediators able to provoke attacks of acute respiratory illness in susceptible individuals.[26] Other potential mechanisms include impairment of lung defences and physiological disturbances of gas transfer. Fine particles do penetrate into the respiratory region, and recent studies have shown that exposure to 288 μg/m^3 of fine urban particles for six hours a day for three days resulted in 37% mortality in bronchitic rats compared with 0% in control rats.[27] Bronchoconstriction was found in the exposed bronchitic rats, and inflammatory cytokines were detected in both the lung and the heart.

Sulphur dioxide, on the other hand, is a highly reactive gas with short half life indoors. It is a known respiratory irritant and bronchoconstrictor, but its effects seem limited to patients with asthma and bronchitis, although sensitivity to exposure varies widely.[17 28] Exposure to sulphur dioxide may therefore not completely explain the observed increase in mortality; it may rather serve as a surrogate of other substances. Since sulphur dioxide is highly correlated with the levels of fine particles in some American cities, Schwartz *et al* postulated that sulphur dioxide may be a marker of fine particles.[29] However, the fact that we found a consistently significant effect of sulphur dioxide on mortality in all western European cities, whatever the level and composition of particles in each one, may suggest that sulphur dioxide has a direct effect. The role of outdoor peak exposures to sulphur dioxide in the increase of daily mortality should be further investigated.

Geographical differences

An intriguing finding was the difference observed in the effects of black smoke and sulphur dioxide on mortality in western and central eastern European cities. The estimates for the central eastern European cities were either similar to or lower than the ones estimated for sulphur dioxide in Erfurt, former East Germany, over the same concentration range (<200 μg/m^3).[30] The variation between east and west may have been because the pollution measurements were unrepresentative of the population exposure, differences existed in the health of the population (smaller proportion of sensitive individuals), or cities had a different pollutant toxicity or mix, possibly because of the sources of pollutants. Furthermore, the model for seasonal control may fit the data less well in the central eastern cities because of a higher and more variable rate of respiratory illness.

In conclusion, the consistency of results from the western European cities with wide differences in topography, climate, environment, and air pollution sources supports a causal association between exposure to particulate matter and sulphur dioxide and mortality from all causes. The reported relative risks are small, but the short term effects of air pollution are not a trivial public health problem if the ubiquity of air pollution exposure is taken into account.

Key messages

- Evidence is accumulating that air pollution below the levels of national and international standards has adverse short term health effects
- In this study data from 12 European cities showed that increases in sulphur dioxide and particulate matter are associated with increased total mortality
- The effects of the two pollutants seem to be independent
- Associations were stronger and more consistent in western European cities
- Current low levels of sulphur dioxide and particles still affect health and further reductions in pollution are needed

The APHEA collaborative group consists of K Katsouyanni, G Touloumi, E Samoli (Greece, coordinating centre); D Zmirou,

P Ritter, T Barumandzadeh, F Balducci, G Laham (Lyons, France); H E Wichmann, C Spix (Germany); J Sunyer, J Castellsague, M Saez, A Tobias (Spain); J P Schouten, J M Vonk, A C M de Graaf (Netherlands); A Ponka (Finland); H R Anderson, A Ponce de Leon, R Atkinson, J Bower, D Strachan, M Bland (UK); W Dab, P Quenel, S Medina, A Le Tertre, B Thelot, B Festy, Y Le Moullec, C Monteil (Paris); B Wotjyniak, T Piekarski (Poland); M A Vigotti, G Rossi, L Bisanti, F Repetto, A Zanobetti (Italy); L Bacharova, K Fandakova (Slovakia).

Funding: The APHEA project was supported by the European Commission, DGXII, Environment 1991-94 Programme (contract number EV5V CT92-0202).

Conflict of interest: None.

1 Ware JH, Tribodeau LA, Speizer FE, Colome S, Ferris BG. Assessment of the health effects of atmospheric sulfur oxides and particulate matter: evidence from observational studies. *Environ Health Perspect* 1981;42:255-76.
2 World Health Organisation Regional Office for Europe. *Air quality guidelines for Europe*. Copenhagen: WHO Regional Publications, 1987. (European series No 23.)
3 Schwartz J. Air pollution and daily mortality: a review and meta-analysis. *Environ Res* 1994;64:36-52.
4 Dockery DW, Pope CA III. Acute respiratory effects of particulate air pollution. *Annu Rev Public Health* 1994;15:107-32.
5 Utell MJ, Samet JM. Particulate air pollution and health. New evidence on an old problem. *Am Rev Respir Dis* 1993;147:1334-5.
6 Hatzakis A, Katsouyanni K, Kalandidi A, Day N, Trichopoulos D. Short term effects of air pollution on mortality in Athens. *Int J Epidemiol* 1986;15:73-81.
7 Derrienic F, Richardson S, Mollie A, Lellouch J. Short term effects of sulphur dioxide pollution on mortality in two French cities. *Int J Epidemiol* 1989;18:186-97.
8 Wichmann HE, Mueller W, Allhoff P, Beckmann M, Bocter N, Csicsaky MJ, *et al*. Health effects during a smog episode in West Germany in 1985. *Environ Health Perspect* 1989;79:89-99.
9 Pönkä A. Asthma and low level air pollution in Helsinki. *Arch Environ Health* 1991;46:262-70.
10 Sunyer J, Saez M, Murillo C, Castellsague J, Martinez F, Anto JM. Air pollution and emergency room admissions for chronic obstructive pulmonary disease: a 5-year study. *Am J Epidemiol* 1993;137:701-5.
11 Katsouyanni K, Zmirou D, Spix C, Sunyer J, Schouten JP, Pönkä A, *et al*. Short-term effects of air pollution on health: a European approach using epidemiological time-series data. *Eur Respir J* 1995;8:1030-8.
12 Katsouyanni K, Schwartz J, Spix C, Touloumi G, Zmirou D, Zanobetti A, *et al*. Short-term effects of air pollution on health: a European approach using epidemiologic time series data. The APHEA protocol. *J Epidemiol Community Health* 1996;50:S12-8.
13 Schwartz J, Spix C, Touloumi G, Bacharova L, Barumandzadeh T, Le Tertre A, *et al*. Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol Community Health* 1996;50:S3-11.
14 The APHEA project. Short-term effects of air pollution on health: a European approach using epidemiological time series data. *J Epidemiol Community Health* 1996;50(suppl):S19-80.
15 Anderson HR, Ponce de Leon A, Bland JM, Bower J, Strachan DP. Air pollution and daily mortality in London: 1987-1992. *BMJ* 1996;312:665-9.
16 Der Simonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177-88.
17 Samet JM, Zeger SL, Berhane K. The association of mortality and particulate air pollution. In: *Particulate air pollution and daily mortality. The phase I report of the Particle Epidemiology Evaluation Project*. Boston, US: Health Effects Institute, 1995.
18 Kinney PL, Ito K, Thurston GD. A sensitivity analysis of mortality/ PM_{10} associations in Los Angeles. *Inhal Toxicol* 1995;7:59-69.
19 Verhoeff AP, Hoek G, Schwartz J, Van Wijnen JH. Air pollution and daily-mortality in Amsterdam. *Epidemiology* 1996;7:225-30.
20 Touloumi G, Pocock J, Katsouyanni K, Trichopoulos D. Short-term effects of air pollution on daily mortality in Athens: a time-series analysis. *Int J Epidemiol* 1994;23:957-67.
21 Wietlisbach V, Pope CA, Ackermann-Liebrich U. Air pollution and daily mortality in three Swiss urban areas. *Soz Praventivmed* 1996;41:107-15.
22 Ballester F, Corella D, Perez-Hoyos S, Hervas A. Air pollution and mortality in Valencia (Spain): a study using the APHEA methodology. *J Epidemiol Community Health* 1996;50:527-33.
23 Suh HH, Spengler JD, Koutrakis P. Personal exposures to acid aerosols and ammonia. *Environ Sci Technol* 1992;26:2507-17.
24 Touloumi G, Katsouyanni K, Zmirou D, Schwartz J, Spix C, Ponce de Leon A, *et al*. Short-term effects of ambient oxidants exposure on mortality: a combined analysis within the APHEA project. 1997 *Am J Epidemiol* (in press.)
25 Department of Health Committee on the Medical Effects of Air Pollution. *Non-biological particles and health*. London: HMSO, 1995.
26 Seaton A, McNee W, Donaldson K, Godden D. Particulate air pollution and acute health effects. *Lancet* 1995;345:176-8.
27 Godleski JJ, Sioutas C, Katler M, Koutrakis P. Death from inhalation of concentrated ambient air particles in animal models. *Am J Respir Crit Care Med* 1996;153:A15.
28 Department of Health Advisory Group on the Medical Effects of Air Pollution Episodes. *Sulphur dioxide, acid aerosols, and particles*. London: HMSO, 1995.
29 Schwartz J, Dockery DW, Neas CM, Wypij D, Ware JH, Spengler JD, *et al*. Acute effects of summer air pollution on respiratory symptom reporting in children. *Am J Respir Crit Care Med* 1994;150:1234-42.
30 Spix C, Heinrich J, Dockery D, Schwartz J, Volksch G, Schwinkowski K, *et al*. Air pollution and daily mortality in Erfurt, East Germany, 1980-1989. *Environ Health Perspect* 1993;101:518-26.

(Accepted 13 March 1997)

Simultaneous immunisation with influenza vaccine and pneumococcal polysaccharide vaccine in patients with chronic respiratory disease

T J Fletcher, W S Tunnicliffe, K Hammond, K Roberts, J G Ayres

Pneumococcal disease is an important cause of morbidity and mortality in the United Kingdom.[1] The increasing numbers of elderly people and the development of drug resistant *Streptococcus pneumoniae* will exacerbate this problem. The safety, efficacy, and cost effectiveness of immunisation with pneumococcal polysaccharide vaccine is established,[2] and immunisation is now recommended for all patients aged over 2 years with chronic lung disease, a target group for influenza vaccination.[3] Coadministration of the vaccines is recommended, though the efficacy of pneumococcal polysaccharide vaccine given in this way has been questioned.[4] We examined whether an immunoresponsive interaction exists between 23 valent pneumococcal polysaccharide vaccine (Pnu-Imune 23) and influenza vaccine (Fluarix).

Patients, methods, and results

One hundred and fifty two adults with chronic respiratory disease were randomised to receive either pneumococcal vaccination and influenza vaccination on the same day (concurrent group; n=76) or influenza vaccination first, followed by pneumococcal vaccination one month later (interval group; n=76). The pneumococcal vaccine was given into the left deltoid muscle and the influenza vaccine into the right deltoid muscle. At the initial visit and one month after each injection venesection was performed for blinded analysis of pneumococcal antibody titres (serotypes 4, 6B, 14, 18C, 19F, 23F) by enzyme linked immunosorbent assay (ELISA) and influenza antibody titres (strains A (Taiwan), A (Johannesburg), B (Harbin)) by

Chest Research Institute, Birmingham Heartlands Hospital, Birmingham B9 5SS
T J Fletcher, *research fellow*
W S Tunnicliffe, *research fellow*
K Hammond, *research nurse*
K Roberts, *research scientist*
J G Ayres, *professor of respiratory medicine*
Correspondence to: Professor Ayres.

BMJ 1997;314:1663–5

Table 1 Demography and serological responses to pneumococcal and influenza vaccination by treatment group

	Concurrent group	Interval group	Intergroup comparison	P value
Demography				
No recruited	76	76		
No completed	73	58		0.001†
Median age (years)	47.9	49.9		0.48‡
No (%) male	36 (49.3)	22 (37.9)		
No (%) female	37 (50.7)	36 (62.1)		
Mean % of predicted forced expiratory volume in one second	77.7	79.2		0.87§
Pneumococcal serotype				
Serotype 4:				
Geometric mean titre before vaccination (mg/l) (95% CI)	0.61 (0.45 to 0.86)	0.63 (0.45 to 0.83)		
Geometric mean titre after vaccination (mg/l) (95% CI)	2.05 (1.46 to 2.89)	2.51 (1.85 to 3.41)		
Geometric mean ratio (95% CI)	3.35 (2.61 to 4.29)	4.01 (2.96 to 5.45)		
Concurrent group to interval group geometric mean ratio (95% CI)			0.84 (0.57 to 1.23)	0.35¶
Proportion (%) with >2-fold rise	46/73 (63.0)	42/58 (72.4)		0.27†
Serotype 6B:				
Geometric mean titre before vaccination (mg/l) (95% CI)	2.53 (1.32 to 2.86)	1.94 (1.32 to 2.86)		
Geometric mean titre after vaccination (mg/l) (95% CI)	8.31 (6.15 to 11.24)	7.35 (4.95 to 10.89)		
Geometric mean ratio (95% CI)	3.28 (2.60 to 4.15)	3.79 (2.75 to 5.20)		
Concurrent group to interval group geometric mean ratio (95% CI)			0.87 (0.59 to 1.27)	0.46¶
Proportion (%) with >2-fold rise	45/73 (61.6)	41/58 (70.7)		0.36†
Serotype 9V:				
Geometric mean titre before vaccination (mg/l) (95% CI)	1.43 (1.09 to 1.88)	1.10 (0.78 to 1.45)		
Geometric mean titre after vaccination (mg/l) (95% CI)	5.90 (4.41 to 7.88)	4.78 (3.54 to 6.45)		
Geometric mean ratio (95% CI)	4.13 (3.23 to 5.28)	4.36 (3.35 to 5.66)		
Concurrent group to interval group geometric mean ratio (95% CI)			0.95 (0.66 to 1.36)	0.78¶
Proportion (%) with >2-fold rise	53/73 (72.6)	44/58 (75.9)		0.69†
Serotype 14:				
Geometric mean titre before vaccination (mg/l) (95% CI)	1.66 (1.13 to 2.44)	2.27 (1.48 to 3.51)		
Geometric mean titre after vaccination (mg/l) (95% CI)	9.22 (5.98 to 14.22)	17.87 (11.40 to 28.03)		
Geometric mean ratio (95% CI)	5.56 (3.85 to 8.02)	7.86 (5.59 to 11.05)		
Concurrent group to interval group geometric mean ratio (95% CI)			0.71 (0.43 to 1.17)	0.18¶
Proportion (%) with >2-fold rise	55/73 (75.3)	51/58 (87.9)		0.08†
Serotype 18C:				
Geometric mean titre before vaccination (mg/l) (95% CI)	1.65 (1.28 to 2.12)	1.57 (1.14 to 2.17)		
Geometric mean titre after vaccination (mg/l) (95% CI)	7.95 (5.89 to 10.73)	6.81 (4.81 to 9.63)		
Geometric mean ratio (95% CI)	4.82 (3.82 to 6.08)	4.34 (3.17 to 5.94)		
Concurrent group to interval group geometric mean ratio (95% CI)			1.11 (0.76 to 1.62)	0.5¶
Proportion (%) with >2-fold rise	56/73 (76.7)	44/58 (75.9)		0.99†
Serotype 19F:				
Geometric mean titre before vaccination (mg/l) (95% CI)	2.7 (1.96 to 3.73)	2.44 (1.63 to 3.63)		
Geometric mean titre after vaccination (mg/l) (95% CI)	7.71 (5.26 to 11.31)	6.16 (4.13 to 9.19)		
Geometric mean ratio (95% CI)	2.85 (2.27 to 3.59)	2.53 (1.93 to 3.31)		0.5¶
Concurrent group to interval group geometric mean ratio (95% CI)			1.13 (0.80 to 1.62)	
Proportion (%) with >2-fold rise	44/73 (60.3)	32/58 (55.2)		0.60†
Serotype 23F:				
Geometric mean titre before vaccination (mg/l) (95% CI)	1.19 (0.88 to 1.62)	1.27 (0.86 to 1.19)		
Geometric mean titre after vaccination (mg/l) (95% CI)	5.11 (3.76 to 6.96)	5.04 (3.36 to 7.56)		
Geometric mean ratio (95% CI)	4.28 (3.33 to 5.51)	3.98 (2.92 to 5.41)		
Concurrent group to interval group geometric mean ratio (95% CI)			1.08 (0.73 to 1.59)	0.7¶
Proportion (%) with >2-fold rise	55/73 (75.3)	43/58 (74.1)		0.99†
Influenza strain				
Strain A [Taiwan]:				
Geometric mean titre before vaccination (mg/l) (95% CI)	49.53 (31.95 to 76.77)	52.05 (32.08 to 84.44)		
Geometric mean titre after vaccination (mg/l) (95% CI)	415.48 (316.62 to 545.21)	314.79 (207.78 to 476.91)		
Geometric mean ratio (95% CI)	8.39 (5.71 to 12.33)	6.05 (4.09 to 8.95)		
Concurrent group to interval group geometric mean ratio (95% CI)			1.39 (0.80 to 2.40)	0.24¶
Proportion (%) with >4-fold rise	49/73 (67.1)	38/57 (66.7)††		0.99†
Strain A [Johannesburg]:				
Geometric mean titre before vaccination (mg/l) (95% CI)	16.15 (11.26 to 23.18)	16.59 (10.41 to 26.46)		
Geometric mean titre after vaccination (mg/l) (95% CI)	181.88 (123.76 to 267.30)	205.67 (134.75 to 313.94)		
Geometric mean ratio (95% CI)	11.26 (8.31 to 15.26)	12.39 (7.86 to 19.54)		
Concurrent group to interval group geometric mean ratio (95% CI)			0.91 (0.54 to 1.53)	0.72¶
Proportion (%) with >4-fold rise	59/73 (80.6)	45/57 (78.9)††		0.83‡
Strain B [Harbin]:				
Geometric mean titre before vaccination (mg/l) (95% CI)	5.22 (3.89 to 7.00)	4.20 (3.18 to 5.54)		
Geometric mean titre after vaccination (mg/l) (95% CI)	24.3 (17.51 to 33.71)	22.49 (15.27 to 33.12)		
Geometric mean ratio (95% CI)	4.66 (3.52 to 6.16)	5.36 (3.83 to 7.49)		
Concurrent group to interval group geometric mean ratio (95% CI)			0.87 (0.57 to 1.33)	0.54¶
Proportion (%) with >4-fold rise	43/73 (58.9)	37/57 (64.9)††		0.59†

†χ^2 Test. ‡Wilcoxon test. §Student's *t* test. ¶Analysis of variance. ††Baseline influenza serological data were unavailable for one subject in the interval group, denominator 57.

haemagglutination inhibition. Patients recorded local and systemic side effects for four days after each vaccination.

Three patients in the concurrent group and 18 in the interval group failed to complete the study. Geometric mean titres, geometric mean ratios (geometric mean titres after vaccination/geometric mean titres before vaccination), and the proportions of patients in each group with a greater than twofold rise in pneumococcal antibody titres and greater than fourfold rise in influenza antibody titres were compared (table 1).

There were no significant differences in serological responses between the groups. The incidence and severity of both local side effects (pain, redness, swelling) and systemic side effects were also similar in the treatment groups. Mild local reactions occurred in 49 subjects, 28 (38%) in the concurrent group and 21 (36%) in the interval group. Systemic reactions occurred in three and five patients in the respective groups.

Comment

Patients with chronic lung disease should be offered pneumococcal and yearly influenza vaccination.[3] In the United Kingdom targeting these patients in general practice has resulted in reasonable influenza vaccination rates but uptake of pneumococcal vaccination in this group has been poor.[5] Our findings suggest that in adults with chronic respiratory disease concurrent immunisation with 23 valent pneumococcal polysaccharide vaccine and influenza vaccine is as well tolerated and immunologically effective as interval vaccination. Our sixfold greater drop out rate in the interval group as compared with the concurrent group raises the possibility that bias could explain our findings. However, there was no significant difference in clinical characteristics between patients who completed the study and those who did not. The drop out rate in the interval group illustrates the practical difficulty of getting patients to return for repeated injections.

With the skills and infrastructure in place for influenza vaccination greater coverage of the population at risk from pneumococcal disease could readily be achieved. When appropriate the opportunity to offer and administer pneumococcal vaccination to adults with respiratory disease at the time of their influenza vaccination should not be missed.

Funding: Wyeth Lederle Vaccines and Paediatrics.
Conflict of interest: None.

1 Macfarlane J. The clinical impact of pneumococcal disease. In: Mayon-White RT, ed. *The clinical impact of pneumococcal disease and strategies for its prevention.* London: Royal Society of Medicine Press, 1995:9-16.
2 Butler I, Breiman RF, Campbell JF, Lipman HB, Broome CV, Facklam RR. Pneumococcal polysaccharide vaccine efficacy. An evaluation of current recommendations. *JAMA* 1993;270:1826-31.
3 Salisbury DM, Begg NT, eds. *Immunisation against infectious diseases, 1992.* London: HMSO, 1992.
4 Huang K-L, Ruben FL, Rinaldo CR, Kingsley L, Lyter DW, Ho M. Antibody responses after influenza and pneumococcal immunization in HIV-infected homosexual men. *JAMA* 1987;257:2047-50.
5 Fedson DS. A policy for the prevention of pneumococcal disease. In: Mayon-White RT, ed. *The clinical impact of pneumococcal disease and strategies for its prevention.* London: Royal Society of Medicine Press, 1995:49-61.

(Accepted 6 December 1997)

Drug points

Acute dissection of the aorta with amphetamine misuse

W C Dihmis, P Ridley, J P Dhasmana, J D Wisheart, Department of Cardiac Surgery, Bristol Royal Infirmary, Bristol BS2 8HW

A 27 year old man who was an intravenous amphetamine misuser presented with a cold ischaemic left leg and chest pain after cutting wood. Dissection of the entire aorta was seen on computed tomography. He was transferred to this hospital, where transthoracic echocardiography confirmed the diagnosis. Surgery showed an intimal tear in the ascending aorta. The segment of aorta containing the intimal tear was excised, the aortic wall reconstituted, and continuity restored with an interposition graft. He died 32 days later as a result of septicaemia and multiorgan failure. Analysis of the blood samples taken at presentation showed a plasma amphetamine concentration of 0.14 mg/l.

Acute dissection of the aorta is rare in young adults.[1 2] The use of amphetamine enhances noradrenaline release, causing surges in blood pressure. Misuse of cocaine, which has a similar action, has been implicated in the pathogenesis of aortic dissection.[3 4] To our knowledge, there have been no reports associating acute dissection of the aorta with amphetamine use. Our patient was chopping wood, a strenuous exercise with an isometric component, which may have contributed to a surge in blood pressure. Serum amphetamine concentrations of 0.5 mg/l or less are not usually associated with serious toxicity, although the half life for amphetamine is short and thus the measured plasma concentration of 0.14 mg/l indicates that amphetamine was present and that the concentration had been higher.

The amphetamines taken by this patient may have been contaminated with other drugs with an additive effect.

1 Larson EW, Edwards WD. Risk factors for aortic dissection: a necropsy study of 161 cases. *Am J Cardiol* 1984;53:849.
2 Wilson SK, Hutchins GM. Aortic dissecting aneurysms: causative factors in 204 subjects. *Arch Pathol Lab Med* 1982;206:175.
3 Barth CW III, Bray M, Roberts WC. Rupture of the ascending aorta during cocaine intoxication. *Am J Cardiol* 1986;57:496.
4 Om A, Porter T, Mohanty PK. Trans-oesophageal echocardiogram diagnosis of acute aortic dissection complicating cocaine abuse. *Am Heart J* 1992;123:532-4.