Papers

Incidence of disorders of spermatogenesis in middle aged Finnish men, 1981-91: two necropsy series

Jarkko Pajarinen, Pekka Laippala, Antti Penttila, Pekka J Karhunen

Abstract

Objective: To investigate if the incidence of disorders of spermatogenesis and testicular tissue morphology have changed in middle aged Finnish men over 10 years.
Design: Two necropsy series completed in 1981 and in 1991.
Setting: Department of Forensic Medicine, University of Helsinki, Finland.
Subjects: 528 men, aged 35 to 69 years, subjected to medicolegal necropsy.
Main outcome measures: Scoring of spermatogenesis and morphometric analysis of testicular tissue components. Individual risk factors for testicular disorders obtained by postmortem blind interviews with acquaintances.
Results: Normal spermatogenesis was found in 41.7% of the men (mean age 53.1 years). Between 1981 and 1991, the ratio of normal spermatogenesis decreased significantly (odds ratio 3.5; 95% confidence interval 2.5 to 5.1) from 56.4% to 26.9%, with a parallel increase in the incidence of partial and complete spermatogenic arrest (2.1; 1.4 to 2.9 and 2.9; 1.7 to 5.0, respectively). During this period, the size of seminiferous tubules decreased, the amount of fibrotic tissue increased, and the weight of testicles decreased significantly. Alterations in testicular characteristics over time could not be explained by changes in body mass index, smoking, alcohol drinking, or exposure to drugs.
Conclusions: The incidence of normal spermatogenesis decreased among middle aged Finnish men from 1981 to 1991, and the incidence of disorders of spermatogenesis and pathological alterations in testicles increased. Deteriorating spermatogenesis may thus be one important factor in the explanation of declining sperm counts observed worldwide.

Introduction

In 1992 Carlsen *et al* reported a significant decrease in the quality of human semen during the past 50 years, including a deterioration in mean semen volume and mean sperm concentration in semen from voluntary sperm donors.[1] Although contradictory and critical results have been published,[2-5] this observation has subsequently been corroborated by several reports additionally suggesting that a similar deterioration may have taken place in the morphology and motility of sperm.[6-8] The reasons for the declining quality of semen and sperm are subjects of current research. Several environmental toxins and chemicals such as alcohol,[9] drugs,[10] industrial solvents,[11] and endogenous and exogenous oestrogen-like compounds[12] have been suggested to alter the function of male reproductive organs via multiple mechanisms, including endocrinological disorders and direct toxic damages on gonadal cells.[13 14]

Thus far, most reports suggesting a deterioration of human semen and sperm during the past decades have analysed the quality of semen from voluntary sperm donors and men attending infertility clinics. Nothing is known, however, about the changes in quality of sperm from middle aged or older men, who may have been exposed to environmental and other risk factors for a longer time than those younger subjects in previous studies. In addition to high selection of subjects in previous studies, which possibly biases the outcome, semen analysis may not relate directly to the severity of testicular lesion at the level of spermatogenesis,[15 16] although it reflects well the overall function of the male reproductive tract. Therefore, to what extent changes in semen quality can be explained by a deterioration of spermatogenesis or some other disorder(s) in male sex organs is undetermined. Alterations in semen quality of donors may even be unrelated to spermatogenesis, reflecting other factors such as changes in the frequency of coitus, or be due to non-toxic exogenous means such as variations in scrotal temperature because, for example, differences in types of clothing. Moreover, it is not known if changes in some known individual risk factors might be associated with declining sperm counts. In this study, we investigated possible changes in incidence of disorders of spermatogenesis and testicular morphology in two necropsy series between 1981 and 1991, comprising middle aged men with no selection as regards spermatogenesis and fertility. Data on exposure to risk factors for testicular disorders were obtained by interviewing close relatives or friends of the deceased.

Subjects and methods

Necropsy series

We examined two necropsy series comprising middle aged men who were subjected to medicolegal necropsy at the Department of Forensic Medicine, University of

Department of Forensic Medicine, University of Helsinki, PO Box 40, 00014, Finland
Jarkko Pajarinen, *doctor*
Antti Penttila, *professor*
Department of Biometrics, University of Tampere, Finland
Pekka Laippala, *professor*
School of Medicine, University of Tampere, Finland
Pekka J Karhunen, *professor*
Correspondence to: Dr Pajarinen.

BMJ 1997;314:13–8

Table 1 Features of necropsy series. Figures are means (SD; 95% confidence intervals) unless stated otherwise

	1981(n=264)	1991(n=264)	Difference*
Age (years)	54.0 (9.6; 52.8 to 55.1)	52.2 (9.6; 51.1 to 53.2)	1.8 (0.2 to 3.4)†
Body mass index (kg/m^2)	24.1 (4.4; 23.6 to 24.6)	25.8 (4.6; 25.2 to 26.4)	1.7 (0.9 to 2.5)†
No (%) according to cause of death:			
Cardiovascular disease	138 (52)	126 (48)	4.6 (–4.0 to 13.1)
Other disease	40 (15)	39 (15)	0.4 (–5.7 to 6.5)
Intoxication	23 (9)	31 (12)	3.0 (–2.1 to 8.2)
Other violent cause	61 (23)	63 (24)	0.8 (–6.5 to 8.0)
Unknown	2 (1)	5 (2)	1.1 (–0.8 to 3.1)
Interval between death and necropsy (days)	3.5 (1.6; 3.3 to 3.7)	3.8 (2.1; 3.6 to 4.1)	0.3 (–0.0 to 0.6)

*Difference in means or percentages (95% confidence intervals) between series. 95% Confidence intervals excluding 0 for means and percentages are considered to be significant.
†Significant difference in comparison between series.

Helsinki. The capital Helsinki and the surrounding Uusimaa province in southern Finland are the most densely populated parts of Finland, with nearly one fifth (1 million) of the population of the country. About 2500 medicolegal necropsies are performed annually in this area. The first series was collected in 1981 and the second in 1991, both as substudies of a study of changes in the risk factors for sudden and violent death. Both series initially comprised 264 consecutive men aged 35 to 69 years (mean 53.1). More than half of the men (65%) died from disease (table 1), cardiovascular diseases being the most common. One third (33.7%) died violently or due to intoxication, most of them accidentally or by committing suicide. In seven cases (1.3%) the cause of death remained unknown. By covering 42% of all deaths of men aged 65 years or less in this area, this series is the most representative sample of middle aged men that can be obtained.

Methods

Testicles were weighed at necropsy, and the middle section of one testis per man was fixed in 10% buffered formalin solution or in Bouin's solution. Histological sections (5 μm) were prepared after samples were dehydrated and embedded in paraffin and were then visualised by Herovic staining.

Spermatogenesis score

In each slide spermatogenesis was scored in 25 randomly chosen cross sections of seminiferous tubules by one member of the study group (JP) with a light microscope at a magnification of × 200. The analysis was performed without knowledge of interview data or other features and series of the subjects. Each sample was allocated to one of the following groups.

Normal spermatogenesis—Germinal epithelium was normal in most tubules. Sertoli cells, spermatogonia, spermatocytes, spermatids, and spermatozoa were all apparent in the tubules with the epithelium appearing thick and condensed. Men showing modest hypospermatogenesis, characterised by a thin germinal epithelium with diffusely reduced cell quantity at all stages of spermatogenesis, were included in this group.

Spermatogenic arrest—All tubules contained at least Sertoli cells and spermatogonia. No mature spermatozoa were observed, suggesting an arrest in the normal spermatogenesis. Complete spermatogenic arrest was defined as a state in which all tubules exhibited arrest of spermatogenesis, whereas for partial spermatogenic arrest some of the seminiferous tubules ($> 5\%$) showed normal spermatogenesis.

The Sertoli cell only syndrome—A complete absence of germinal cells was observed. Only Sertoli cells remained in most tubules, and in some areas tubules were obliterated by fibrosis. In some cases with complete spermatogenic arrest some of the tubules showed only Sertoli cells, but these cases were grouped on the basis of the main features of the section.

Fibrotic tubules—All seminiferous tubules were totally obliterated by fibrotic tissue with both germinal cells and Sertoli cells missing.

Morphometric analysis of testicular tissue components

Measurement of the percentage area covered by fibrotic tissue and Leydig cells was performed with a light microscope equipped with an ocular grid with 100 defined elements. The number of these elements situated on each of the testicular tissue components was then calculated for one of three randomly chosen cross sections of testicular tissue from each man at a magnification of × 40. An average percentage value for tissue components was calculated. Mean diameters of seminiferous tubules were measured in a proportioned segment of line on the microscope view on a computer screen. Ten randomly chosen circular cross sections of seminiferous tubules were measured, and an average diameter was calculated.

Scoring of spermatogenesis, morphometric analysis of areas of tissue components of testis, and measurement of the diameter of seminiferous tubules were performed blind to interview data.

Table 2 Details of risk factors for testicular disorders obtained through interviews with acquaintances in two necropsy series. Figures are numbers (percentages) of subjects

Series	1981 (n=264)	1991 (n=264)	Difference*
Acquaintances interviewed	151 (57)	109 (41)	15.9 (7.5 to 24.3)†
Use of drugs:			
Antihypertensive drugs	15 (10)	5 (5)	5.4 (–0.8 to 11.5)
Sedatives or tranquillisers	68 (45)	54 (50)	4.5 (–7.8 to 16.8)
Other	42 (28)	32 (29)	1.5 (–9.6 to 12.7)
Smoking	103 (68)	81 (74)	6.1 (–5.0 to 17.2)
Use of alcohol:			
Low or moderate drinking	80 (53)	52 (48)	5.3 (–7.0 to 17.6)
Heavy drinking	71 (47)	57 (52)	5.3 (–7.0 to 17.6)

*Difference in means or percentages (95% confidence interval) between series. 95% Confidence intervals excluding 0 for means and percentages considered to be significant.
†Significant difference in comparison between series.

Interview study and risk factors for testicular disorders

To evaluate the role of various individual risk factors for altered spermatogenesis a relative or a close acquaintance of the subject was contacted and interviewed (table 2). The interview comprised a structured set of more than 50 detailed questions on occupation, use of medications, and smoking and drinking habits.[17]

A complete forensic toxicological examination was performed for all men, including a determination of blood and urine alcohol concentration at the time of necropsy and analysis of the presence of drugs and chemicals in blood, stomach, and liver.

Statistical analysis

The data on spermatogenesis were analysed in cross tabulated form by using odds ratios with 95% confidence intervals (CIA software). Confidence intervals excluding the value of 1 were considered to be significant. When analysing testicular features we used two way analysis of covariance. We used covariate structure and adjustment of means to eliminate the confounding effect, when necessary. Post-hoc comparisons were made with Sheffe's test. The data were analysed in logarithmic and square root transformed form, but the results were identical with those based on crude data. The analysis was carried out with Statistica for Windows (version 5.0) on a 486 PC.

Results

Necropsy series

There were no significant differences in the causes of death between the two series. The interval between death and necropsy was slightly longer in the 1991 series, in which the men were also slightly younger and had a significantly higher mean body mass index compared with those in the 1981 series. No significant differences were observed between the occupations of men in 1981 and 1991.

Changes in spermatogenesis score and testicular morphometric characteristics

Normal spermatogenesis was found in 41.7% (220) of all the 528 men. In 1981, 56.4% (148) of the men showed normal spermatogenesis, whereas in 1991 the incidence had decreased by more than a half (29.5% percentage points) to 26.9% (71) (odds ratio 3.5; 2.5 to 5.1; table 3). Simultaneously, the incidence of partial spermatogenic arrest increased from 31.4% (83) to 48.5% (128) (2.1; 1.4 to 2.9) and that of complete spermatogenic arrest from 8.0% (21) to 20.1% (53) (2.9; 1.7 to 5.0). There were 15 cases of the Sertoli cell only syndrome and eight cases of fibrotic testicles. No significant change was observed, however, in their incidences between 1981 and 1991.

Testicular weight showed a significant decline from an adjusted mean of 18.9 g in 1981 to 17.8 g in 1991 (table 4). The adjusted mean diameter of seminiferous tubules and morphometrically measured percentage area of seminiferous tubules were also smaller in 1991, with a corresponding increase in the percentage area of fibrotic tissue in the 1991 series.

Risk factors for testicular disorders

Alcohol—The proportion of men with reported moderate and heavy drinking did not change significantly between 1981 and 1991. Disorders of spermatogenesis increased significantly between 1981 and 1991 among both moderate and heavy drinkers (table 5). In both series normal spermatogenesis was less common in heavy drinkers than in moderate drinkers in corresponding series. Testicular weight showed a slight decrease in both consumption groups in both series (table 6).

Drugs—Of the men with interview data available, 122/260 (46.9%) used sedative or tranquillising drugs, whereas 20 (7.7%) were users of antihypertensive medication. The proportion of men using these drugs, however, did not change from 1981 to 1991.

Table 3 Comparison of status of spermatogenesis (numbers (percentages)) between series (1981 *v* 1991) among all subjects

Status of spermatogenesis	1981(n=264)	1991(n=264)	Odds ratio (95% confidence interval)*
Normal	149 (56)	71 (27)	0.3 (0.2 to 0.4)†
Partial arrest	83 (31)	128 (49)	2.1 (1.4 to 2.9)†
Complete arrest	21 (8)	53 (20)	2.9 (1.7 to 5.0)†
Sertoli cell only syndrome	9 (3)	6 (2)	1.5 (0.5 to 4.3)
Fibrosis	2 (1)	6 (2)	3.0 (0.6 to 15.2)

*For difference in comparison between series; 95% confidence intervals excluding 1 are considered to be significant.
†Significant difference in comparison between series.

Smoking—There were 184 (70.8%) reported smokers among the 260 men for whom data were available. Smoking was slightly more common in 1991, but no difference was observed in the status of spermatogenesis between smokers and non-smokers.

Body mass index—There was a significant increase in the mean body mass index of the men between 1981 and 1991. Body mass index, however, showed no statistical association with the occurrence of testicular disorders, although it was a significant covariate for testicular weight, diameter of seminiferous tubules, and testicular morphology.

Pooled risks—Disorders of spermatogenesis were more common in men with pooled risk factors for altered spermatogenesis (heavy drinking and smoking and use of tranquillisers) than in men with none of the mentioned risk factors (odds ratio 3.2). This difference, however, was not significant because of the small numbers of cases in both groups.

Discussion

Declining sperm counts

Carlsen *et al* recently reviewed several reports published from 1938 to 1991 on semen quality of voluntary sperm donors and suggested that sperm counts may have decreased by 42% during the past five decades with a concurrent slight decrease in volume of semen.[1] Although criticism has been directed against statistical methods used in that work[2 3] and controversial reports have been published,[4 18 19] including on the semen quality of Finnish men,[5 20] several results corroborating the observation of Carlsen *et al* have been published in recent years.[6-8] The hypothesis of deteriorating function of the male reproductive tract has thus far been based on findings observed in quality

Table 4 Comparison of status of testicular features between series (1981 *v* 1991) among all subjects. Figures are adjusted means (SD; 95% confidence intervals) unless stated otherwise

Testicular features	1981	1991	Difference*
Testicular weight (g)	18.7 (5.7; 18.0 to 19.4)	18.0 (6.2; 17.2 to 18.7)	0.7 (–0.3 to 1.7)
Adjusted mean	18.9	17.8	1.1 (0.1 to 2.1)†
Seminiferous tubular diameter (μm)	184.9 (25.2; 180.7 to 189.9)	180.9 (26.3; 174.7 to 187.1)	4.0 (–0.4 to 8.4)
Seminiferous epithelium (%)	52.3 (13.1)	47.2 (13.2)	5.1 (2.9 to 7.4)†
Adjusted mean	52.7	47.1	5.6 (3.4 to 7.9)†
Fibrotic tissue (%)	40.8 (12.6; 39.3 to 42.4)	46.8 (13.1; 45.5 to 48.8)	6.0 (3.8 to 8.2)†
Adjusted mean	40.4	47.2	6.8 (4.6 to 9.0)†
Leydig cells (%)	6.9 (3.0; 6.5 to 7.2)	5.7 (3.7; 5.2 to 6.5)	1.2 (0.6 to 1.8)†

*Difference in means (95% confidence interval) between series. 95% Confidence intervals excluding 0 for means considered to be significant.
†Significant difference in comparison between series.

Table 5 Comparison of status of spermatogenesis among moderate and heavy drinkers between series (1981 *v* 1991). Values are numbers (percentages) of subjects

	Moderate drinkers			Heavy drinkers		
Status of spermatogenesis	1981(n=80)	1991(n=52)	Odds ratio (95% confidence interval)*	1981(n=71)	1991(n=57)	Odds ratio (95% confidence interval)*
Normal	62 (78)	16 (31)	0.1 (0.05 to 0.3)†	27 (38)	12 (21)	0.4 (0.2 to 1.0)†
Partial arrest	16 (20)	33 (64)	7.0 (3.2 to 15.3)†	29 (41)	29 (51)	1.5 (0.7 to 3.0)
Complete arrest	2 (3)	3 (6)	2.4 (0.4 to 14.8)	10 (14)	12 (21)	1.6 (0.6 to 4.1)
Sertoli cell only syndrome	—	—	NA	5 (7)	4 (7)	1.0 (0.3 to 3.9)
Fibrosis	—	—	NA	—	—	N/A

NA=not applicable.
*For difference in comparison between series; 95% confidence intervals excluding 1 considered to be significant.
†Significant difference in comparison between series.

of semen and sperm, which reflect well the overall function of male reproductive organs. Meinhard *et al*, however, reported normal spermatogenesis or blockage of seminiferous tubules, a status in which spermatogenesis proceeds normally, in half of 100 infertile men who were oligospermic.[15] Similarly, in a recent study, fine needle cytology indicated normal spermatogenesis in almost one third of 534 azoospermic men.[16]

Thus, semen analysis may, in fact, not be a good guide to the severity of testicular lesion at the level of spermatogenesis, and therefore it is questionable whether or not the declining changes in sperm counts are due to increases in disorders of spermatogenesis or whether other disorders in the male reproductive tract are involved. In the present study, we found that the incidence of normal spermatogenesis has decreased significantly among middle aged men, with a parallel increase in the rate of disorders of spermatogenesis during an interval of 10 years between 1981 and 1991. This finding suggests that the quality and dispatch of spermatogenesis are deteriorating in middle aged men and also confirms earlier presumptions on deteriorating spermatogenesis being the main cause of decreasing sperm counts.

Sperm counts in Finland

Vierula *et al* reported recently that sperm density and total sperm counts of Finnish men have not gone through any changes in the past 28 years,[5] confirming earlier findings on good semen quality of Finnish men.[20] Although we found a significant decrease in the incidence of normal spermatogenesis in the past 10 years, our results do not inevitably disagree with those of Vierula *et al*. By definition, testes expressing partial spermatogenic arrest, a slight and probably reversible disorder which was most commonly found among men in 1991 series, still produce various amounts of mature spermatozoa. Therefore, if semen analysis is performed after a few days of abstinence before sperm donation, men with partial spermatogenic arrest may not differ significantly from those with completely normal spermatogenesis. The subjects in our study, however, were on average 20 years older than those of Vierula *et al*, and thus the results are not fully comparable. In addition, our material came from the densely populated province of Uusimaa, mostly in and around Helsinki, whereas that of Vierula *et al* originated from less densely populated and industrialised communities several hundred kilometres north east of Uusimaa province. Recent work from France has shown a significant decline in sperm counts during an interval of 20 years between 1973 and 1992,[6] possibly suggesting that geographical variations do exist in the abundance and distribution of factors causing disorders of male reproductive tract, as also hypothesised by Vierula *et al*. Two recently published reports also found unchanged sperm counts in the United States over the past few decades, suggesting that deterioration of semen quality is not geographically uniform.[18 19] Also, the fertility of Finnish couples was recently observed to be significantly greater than that of British couples, further supporting this hypothesis.[21]

Disorders of male reproductive tract

In addition to declining sperm counts, the incidences of testicular cancer and specific disorders of male reproductive tract have increased in recent decades.[22-25] Both observations have provoked an active discussion of a possible common aetiology. Several environmental factors are known to cause alterations in the male reproductive tract, such as certain drugs, chemicals,[10 11]

Table 6 Comparison of status of testicular features among moderate and heavy drinkers between series (1981 *v* 1991). Figures are adjusted means (SD; 95% confidence intervals) unless stated otherwise

	Moderate drinkers			Heavy drinkers		
Testicular features	1981(n=80)	1991(n=52)	Difference*	1981(n=71)	1991(n=57)	Difference*
Testicular weight (g)	20.0 (5.9; 18.7 to 21.3)	20.2 (6.0; 18.6 to 21.8)	0.2 (–1.9 to 2.3)	16.9 (4.9; 15.7 to 18.0)	17.2 (4.6; 16.0 to 18.4)	0.3 (–1.4 to 2.0)
Adjusted mean	20.5	19.9	0.6 (–1.5 to 2.7)	17.1	17.0	0.1 (–1.6 to 1.8)
Seminiferous tubular diameter (μm)	190.9 (23.5; 186 to 196)	185 (22; 177 to 193)	5.9 (–2.2 to 14.0)	175 (26; 167 to 184)	176 (30; 165 to 187)	1.0 (–8.8 to 10.8)
Adjusted mean	191.9	181.1	10.8 (2.7 to 18.9)†	176.9	174.9	2.0 (–7.8 to 11.8)
Seminiferous epithelium (%)	55.1 (11.2; 50.7 to 53.9)	50.4 (9.2; 45.4 to 48.9)	4.7 (1.0 to 8.4)†	48.5 (14.5; 45.1 to 51.9)	48.7 (11.0; 45.9 to 51.6)	0.2 (–4.4 to 4.8)
Adjusted mean	56.5	50.3	6.2 (2.5 to 9.9)†	48.6	47.9	0.7 (–3.9 to 5.3)
Fibrotic tissue (%)	38.0 (10.6; 35.7 to 40.3)	43.9 (9.1; 41.5 to 46.4)	5.9 (2.4 to 9.4)†	44.0 (14.2; 40.7 to 47.3)	45.2 (11.3; 42.2 to 48.1)	1.2 (–3.4 to 5.8)
Adjusted mean	36.5	44.2	7.7 (4.2 to 11.2)†	43.9	46.1	2.2 (–2.4 to 6.8)
Leydig cells (%)	6.9 (3.2; 6.2 to 7.6)	5.5 (3.5; 4.7 to 6.6)	1.4 (0.2 to 2.6)	7.5 (2.8; 6.8 to 8.1)	6.1 (3.2; 5.2 to 6.9)	1.4 (0.4 to 2.4)†

*Difference in means (95% confidence interval) between series. 95% Confidence intervals excluding 0 for means considered to be significant.
†Significant difference in comparison between series.

and heavy metals as well as excessive drinking.[26] The role of the increase in environmental oestrogen as a possible common denominator to the adverse effects on male reproductive tracts has achieved much emphasis during recent years.[12-29] The suggested increase of oestrogens or oestrogen-like compounds in past decades may originate from several sources, such as from diet,[30] which has gone through many changes in industrialised countries, or from the increasing use of many organochlorides that act like oestrogens[31] and may accumulate in fat tissue. It has also been hypothesised that fat tissue may convert certain steroids to oestrogens and that increasing body fat content may lead to an increase in the bioavailability of oestrogens through a decrease in the concentration of sex hormone binding globulin. In our work, men in the 1991 series had a significantly higher mean body mass index compared with men 10 years earlier, possibly indicating higher body fat content of men in the latter series. Body mass index, however, showed no association with disorders of spermatogenesis.

The mechanisms of declining sperm counts induced by oestrogen or toxins may entail a disturbance in prepubertal multiplication of Sertoli cells, possibly inducing a decrease in their number. Sertoli cells are involved in the formation of the blood-testis barrier and are known to provide nutritional and mechanical support to spermatogenic cells, thereby sustaining normal spermatogenesis.[32] In the present study, small seminiferous tubules were associated with disorders of spermatogenesis, corroborating our own earlier findings[9] and suggesting a disturbance in the function of Sertoli cells or a decrease in their number as a possible aetiological mechanism for declining sperm counts.

Key messages

- Several recent reports have suggested a significant decrease in human sperm counts over the past few decades
- We used middle aged subjects, with no biasing selection as regards fertility or status of spermatogenesis, to evaluate changes in the incidences of disorders of spermatogenesis from 1981 to 1991
- Normal spermatogenesis was observed significantly less often in the 1991 series than in 1981, whereas the incidences of disorders of spermatogenesis increased significantly during that time
- Changes in the status of spermatogenesis coexisted with decreased testicular weight, smaller seminiferous tubules, and increased fibrosis of testicular tissue
- Alterations in spermatogenesis could not be explained by a change in individual risk factors between the series, such as smoking, drinking, or use of medication, thus challenging further research to illuminate specific reasons for deteriorating spermatogenesis and declining sperm counts

Risk factors for altered spermatogenesis

Excessive use of alcohol is often associated with disorders of spermatogenesis and testicular damage.[33 34] In the present study, we analysed time dependent changes of spermatogenesis over a 10 year period among moderate and heavy drinkers and examined changes in smoking and use of drugs to evaluate the effect of possible changes in the exposure to these individual risk factors. The time differences in spermatogenesis scores were of greater extent among moderate drinkers than among heavy drinkers. This observation is likely to exclude increasing alcohol consumption in Finland between 1981 and 1991[35] as an explanatory factor to widespread deterioration of spermatogenesis. In 1991 moderate drinkers showed incidences of scores of spermatogenesis almost equal to those among heavy drinkers 10 years earlier. Similarly, there was a highly significant increase in the amount of fibrotic tissue in testes of men with moderate drinking habits, whereas no such time dependent change was observed among heavy drinkers. These results suggest that heavy drinkers showing normal spermatogenesis and testicular morphology might be more tolerant to toxic stress originating from the environment. We have also recently reported a possible involvement of a genetic component in the development of alcohol induced disorders of spermatogenesis.[36]

There were more smokers in the 1991 series compared with 1981. A large number of studies have produced controversial results on the effects of smoking on semen and sperm.[37-41] In the present study, no association was observed between smoking and status of spermatogenesis. Additionally, there were no significant changes in the ratio of men who used drugs between 1981 and 1991.

Conclusions

In conclusion, we report a deterioration in male reproductive function at the level of spermatogenesis among middle aged Finnish men between 1981 and 1991. This finding supports several previous observations on declining sperm counts, suggesting that both events might be explained by common genotoxic factors that affect spermatogenesis. Declining spermatogenesis was not explained by different exposure to drugs or incidence of smoking or alcohol consumption, and thus more research is clearly needed to evaluate the roles of different factors for declining sperm counts and deteriorating spermatogenesis.

Funding: Yrjo Jahnsson Foundation, Finnish Foundation for Alcohol Studies, and Medical Society of Finland.

Conflict of interest: None.

1 Carlsen E, Giwercman A, Keiding N. Skakkebaek NE. Evidence for decreasing quality of semen during past 50 years. *BMJ* 1992;305:609-13.
2 Olsen GW, Bodner KM, Ramlow JM, Ross CE, Lipshultz LI. Have sperm countsbeen reduced 50 percent in 50 years? A statistical model revisited. *Fertil Steril* 1995;63:887-93.
3 Bromwich P. Cohen J. Stewart I, Walker. Decline in sperm counts: an artefact ofchanged reference range of "normal." *BMJ* 1994;309:19-22.
4 Bujan L, Mansat A, Pontonnier F, Mieusset R. Time series analysis of sperm concentration in fertile men in Toulouse, France between 1977 and 1992. *BMJ* 1996;312:471-2.
5 Vierula M, Niemi M, Keiski A, Saaranen M, Saarikoski S, Suominen J. High and unchanged sperm counts of Finnish men. *Int J Androl* 1996;19:11-7.
6 Auger J, Kunstmann JM, Czyglik F, Jouannet P. Decline in semen quality among fertile men in Paris during the past 20 years. *N Engl J Med* 1995;332:281-5.
7 Comhair F, Van Waelaghem K, DeClercq N, Vermeulen L, Schoonjans F. Statement on the general reduction in sperm quality. *Int J Androl* 1995;18(suppl 2):1 -2.

8 Irvine S, Cawood E, Richardson D, MacDonald E, Aitken J. Evidence of deteriorating semen quality in the United Kingdom: birth cohort study in 577 men in Scotland over 11 years. *BMJ* 1996;312:467-71.
9 Pajarinen J, Karhunen PJ, Lalu K, Savolainen V, Penttila A, Laippala P. Moderate alcohol consumption and disorders of human spermatogenesis. *Alcohol Clin Exp Res* 1996;20:332-7.
10 Drife JO. The effects of drugs on sperm. *Drugs* 1987;33:610-22.
11 Yamada K. Influence of laquer thinner and some organic solvents on reproductive and accessory reproductive organs in the male rat. *Biol Pharm Bull* 1993;16:425-7.
12 Sharpe, RM, Skakkeback NE. Are oestrogens involved in falling sperm counts and disorders of the male reproductive tract? *Lancet* 1993;341:1392-5.
13 Widenius TV, Orava MM, Vihko RK, Ylikahri RH, Eriksson CJP. Inhibition of testosterone biosynthesis by ethanol: multiple sites and mechanisms in dispersed Leydig cells. *J Steroid Biochem* 1987;28:185-8.
14 Rosenblum E, Gavaler JS, Van Thiel DH. Lipid peroxidation: a mechanism for ethanol-associated testicular injury. *Endocrinology* 1985;115:311-18.
15 Meinhard E, McRae CU, Chisholm GD. Testicular biopsy in evaluation of male infertility. *BMJ* 1973;iii:577-81.
16 Verma AK, Basu D, Jayram G. Testicular cytology in azoospermia. *Diag Cytopathol* 1993;9:37-42.
17 Karhunen PJ, Penttila A. Validity of postmortem alcohol reports. *Alcohol Alcoholism* 1990;25:25-32.
18 Paulsen CA, Bergman NG, Wang CW. Data from men in greater Seattle area reveals no downward trend in semen quality: further evidence that deterioration of semen quality is not geographically uniform. *Fertil Steril* 1996;65:1015-20.
19 Fisch H, Goluboff ET, Olson JH, Feldsburg J, Broder SJ, Barad DH. Semen analysis in 1283 men from the United States over a 25-year period: no decline in quality. *Fertil Steril* 1996;65:1009-14.
20 Suominen J, Vierula M. Semen quality of Finnish men. *BMJ* 1993;306:1579.
21 Joffe M. Decreased fertility in Britain compared with Finland. *Lancet* 1996;347:1519-21.
22 Wanderas EH, Tretli S, Fossa SD. Trends in incidence of testicular cancer in Norway. *Eur J Cancer* 1995;31A:2044-8.
23 Osterlind A. Diverging trends in incidence and mortality of testicular cancer in Denmark, 1943-1982. *Br J Cancer* 1986;53:501-5.
24 Hakulinen T, Andersen AA, Malker B, Pukkala E, Schou G, Tulinus H. Trends in cancer incidence in the Nordic countries. *Acta Pathol Microbiol Immunol Scand* 1986;suppl 288:1-151.
25 Jackson MB and the John Radcliffe Hospital Cryptorchidism Research Group. The epidemiology of cryptorchism. *Horm Res* 1988;30:153-6.
26 Pajarinen JT, Karhunen PJ. Spermatogenic arrest and "Sertoli cell only" syndrome—common alcohol-induced disorders of the human testis. *Int J Androl* 1994;17:292-9.
27 Carlsen E, Giwerchman A, Keiding N. Skakkebaek NE. Declining semen quality and increasing incidence of testicular cancer: is there a common cause? *Environ Health Perspect* 1995;103(suppl 7):137-9.
28 Sharpe RM. Declining sperm counts in men—is there an endocrine cause? *J Endocrinol* 1993;136:357-60.
29 Sharpe RM, Fisher JS, Millar MM, Joblin S, Sumpter JP. Gestational and lactational exposure of rats to xenoestrogens results in reduced testicular size and sperm production. *Environ Health Perspect* 1995;103:1136-43.
30 Adlercreutz H. Diet, breast cancer and sex hormone metabolism. *Ann NY Acad Sci* 1990;595:281-90.
31 Field B, Selub M, Hughes CL. Reproductive effects of environmental agents. *Semin Reprod Endocrinol* 1990;8:44-54.
32 Jegou B. The sertoli cell in vivo and in vitro. *Cell Biol Toxicol* 1992;8:49-54.
33 Boiesen PT, Lindholm J, Hagen C, Bahnsen M, Fabricius-Bjerre N. Histological changes in testicular biopsies from chronic alcoholics with and without liver disease. *Acta Pathol Microbiol Scand* 1979;87:139-42.
34 Kuller LH, May SJ, Perper JA. The relationship between alcohol, liver disease, and testicular pathology. *Am J Epidemiol* 1978;108:192-9.
35 Finnish State Alcohol Company. *Alcohol statistical yearbook*. Helsinki: Finnish State Alcohol Company, 1993.
36 Pajarinen JT, Savolainen VT, Perola M, Penttila A, Karhunen PJ. Glutathione s transterase-M1 "null" genotype and alcohol-induced disorders of human spermatogenesis. *Int J Androl* 1996;19:155-63.
37 Chia SE, Ong CN, Tsakok FM. Effects of cigarette smoking on human semen quality. *Arch Androl* 1994;33:163-8.
38 Holzki G, Gall H, Hermann J. Cigarette smoking and sperm quality. *Andrologia* 1991;23:141-4.
39 Dikshit RK, Buch JG, Mansuri SM. Effect of tobacco consumption on semen quality of a population of hypofertile men. *Fertil Steril* 1987;48:334-6.
40 Evans HJ, Fletcher J, Torrance M, Hargreave TB. Sperm abnormalities and cigarette smoking. *Lancet* 1981;i:627-9.
41 Godfrey B. Sperm morphology in smokers. *Lancet* 1981;i:948.

(Accepted 16 October 1996)

Follow up study of moderate alcohol intake and mortality among middle aged men in Shanghai, China

Jian-Min Yuan, Ronald K Ross, Yu-Tang Gao, Brian E Henderson, Mimi C Yu

Department of Preventive Medicine, USC/ Norris Comprehensive Cancer Center, University of Southern California School of Medicine, Los Angeles, California 90033, USA

Jian-Min Yuan, *research fellow*

Ronald K Ross, *professor*

Mimi C Yu, *professor*

Brian E Henderson, *professor*

BMJ 1997;314:18–23

continued over

Abstract

Objective: To assess the risk of death associated with various patterns of alcohol intake.
Design: Prospective study of mortality in relation to alcohol consumption at recruitment, with active annual follow up.
Setting: Four small, geographically defined communities in Shanghai, China.
Subjects: 18 244 men aged 45-64 years enrolled in a prospective study of diet and cancer during January 1986 to September 1989.
Main outcome measure: All cause mortality.
Results: By 28 February 1995, 1198 deaths (including 498 from cancer, 269 from stroke, and 104 from ischaemic heart disease) had been identified. Compared with lifelong non-drinkers, those who consumed 1-14 drinks a week had a 19% reduction in overall mortality (relative risk 0.81; 95% confidence interval 0.70 to 0.94) after age, level of education, and cigarette smoking were adjusted for. This protective effect was not restricted to any specific type of alcoholic drink. Although light to moderate drinking (28 or fewer drinks per week) was associated with a 36% reduction in death from ischaemic heart disease (0.64; 0.41 to 0.998), it had no effect on death from stroke, which is the leading cause of death in this population. As expected, heavy drinking (29 or more drinks per week) was significantly associated with increased risks of death from cancer of the upper aerodigestive tract, hepatic cirrhosis, and stroke.
Conclusions: Regular consumption of small amounts of alcohol is associated with lower overall mortality including death from ischaemic heart disease in middle aged Chinese men. The type of alcoholic drink does not affect this association.

Introduction

In prospective cohort studies conducted in Western populations men and women who drink light to moderate amounts of alcohol have been found to have lower overall death rates than either non-drinkers or heavy drinkers.[1-11] The observed protective effect of drinking on overall mortality is largely due to a reduced risk of fatal ischaemic heart disease, which accounts for roughly one third of all deaths in such populations. High consumption of wine has been suggested to be responsible for the low risk of ischaemic heart disease in France,[12] and recent findings from a

Department of Epidemiology, Shanghai Cancer Institute, Shanghai 200032, People's Republic of China
Yu-Tang Gao, *professor*

Correspondence to: Dr Yuan.

Table 1 Drinking patterns of Shanghai cohort at recruitment

	Former drinkers (n=383)	No of drinks/ week of current drinkers						All drinkers (n=7773)
		≤7 (n=2354)	8-14 (n=1809)	15-21 (n=1077)	22-28 (n=949)	29-42 (n=712)	≥43 (n=489)	
Mean No of drinks/week	21.4	3.5	10.4	17.3	23.9	34.5	61.7	16.9
Mean age at starting (years)	25.8	32.9	31.0	28.3	28.2	26.0	23.4	29.7
Mean duration (years)	25.6	22.5	24.8	28.1	28.1	30.0	32.8	26.0
Mean daily ethanol intake (g):	39.1	6.3	18.9	31.0	43.7	62.7	112.6	30.7
Beer	2.7	2.1	4.9	3.6	3.9	4.6	7.9	3.8
Wine	8.2	2.9	7.1	19.7	8.6	19.7	26.1	10.2
Spirits	28.3	1.3	6.9	7.8	31.1	38.4	78.6	16.7

prospective study in Copenhagen support this hypothesis; drinking wine but not beer or spirits was found to be associated with reduced overall mortality in Danish people.[13] On the other hand, several studies reported no differential effects of wine and other types of alcoholic beverages on risk of death.[11-17] In a recent review of the relation between specific types of alcoholic drink and risk of coronary heart disease, Rimm *et al* concluded that there is no evidence that wine offers any more protection against ischaemic heart disease than beer or spirits.[18]

We examined prospectively the relation between alcohol consumption and mortality in a large cohort of middle aged Chinese men in Shanghai, China. In this cohort, ischaemic heart disease accounts for only about 9% of all deaths[19] and grape wine is a minor source of alcohol among those who drink regularly. Therefore, our study population is a valuable one to evaluate further the relation between moderate alcohol intake and mortality.

Subjects and methods

Between January 1986 and September 1989 we invited all eligible male residents of four small, geographically defined communities from a wide area of the city of Shanghai to participate in a prospective, epidemiological study of diet and cancer. The men had to be aged 45-64 years and have no history of cancer. At recruitment we interviewed each subject using a structured questionnaire which included level of education, usual occupation, adult height and usual adult weight, history of tobacco and alcohol use, current diet, and medical history.

We asked each participant whether he had ever drunk alcoholic beverages at least once a week for six months or more. If the answer was yes, he was asked to provide the age at which he started to drink regularly and the usual amount of consumption of beer, wine, and spirits separately. If the subject was a former drinker the age at which he stopped drinking was recorded. One drink was defined as 360 g of beer (12.6 g of ethanol), 103 g of wine (12.3 g of ethanol), or 30 g of spirits (12.9 g of ethanol).[20]

Details of the follow up procedures for this cohort have been described.[21] We routinely reviewed death certificates and cancer reports from the population-based Shanghai Cancer Registry, and all surviving members of the cohort were contacted each year.

For each man, years of follow up were counted from the date of recruitment to 28 February 1995 or the date of death or loss to follow up, whichever occurred first. Overall mortality was calculated according to various categories of alcohol consumption. Because of the small number of former drinkers in the study population, former and current drinkers could not be analysed separately. All rates were adjusted for age by using the person-year distribution of the entire cohort as the internal standard.

We used the Poisson regression method to examine the relation between alcohol consumption at recruitment and subsequent risk of death.[22] Relative risks and their corresponding 95% confidence intervals were calculated for various categories of drinkers in reference to lifelong non-drinkers. Two statistical models were used to examine whether the observed relation between alcohol intake and mortality was specific to the type of drink consumed (beer, wine, or spirits). Model one examines the relation between a specific type of alcoholic drink and overall mortality while controlling for the consumption of the other two types (a covariate representing the total ethanol intake for the latter was added to the list of confounding variables). Model two tests for possible association between mortality and a specific type of alcoholic drink after adjusting for total ethanol intake (a covariate representing total ethanol intake was added to the list of confounding variables). Therefore, if the putative U shaped relation between alcohol intake and mortality is independent of type of drink consumed, model one will yield similar U shaped curves for beer, wine, and spirits while model two will show no additional association with beer, wine, or spirits after total alcohol intake has been taken into consideration. The other confounding variables adjusted for in the various regression models were age, cigarette smoking, and level of education.

We used the proportional hazards regression method to test for a curvilinear relation between alcohol intake and total mortality (linear and quadratic terms for number of drinks per week were included in the regression model).[23] All quoted P values are two sided. Relative risks with two sided P values under 0.05 were considered to be significantly different from 1.0.

Results

During the three year recruitment period 18 244 men enrolled in the study, and to date only 108 have been lost to follow up. At recruitment, 57% (10 471) of cohort members had never drunk alcoholic beverages regularly, 41% (7390) were current drinkers, and the remaining 2% (383) were former drinkers (table 1). Drinking was closely and positively associated with cigarette smoking and moderately related to age (posi-

Table 2 Total mortality by alcohol consumption status at recruitment

	Non-drinkers	No of drinks/week of ever drinkers†						All subjects
		≤7	8-14	15-21	22-28	29-42	≥43	
Entire cohort:								
No of deaths	656	136	118	78	74	68	68	1198
Mortality‡	971	826	881	988	1028	1282	1696	980
Relative risk§	1.00	0.80*	0.82*	0.84	0.84	1.03	1.30*	—
Relative risk¶	1.00	0.79*	0.81*	0.83	0.84	1.02	1.28	—
Excluding former drinkers:								
No of deaths	656	120	101	70	67	66	56	1136
Mortality‡	971	772	789	945	984	1317	1529	948
Relative risk§	1.00	0.75**	0.73**	0.80	0.80	1.05	1.16	—
Excluding those with selected serious illness††:								
No of deaths	535	113	96	68	64	59	61	996
Mortality‡	859	732	774	915	947	1194	1633	867
Relative risk§	1.00	0.80*	0.80	0.88	0.86	1.07	1.39*	—
Excluding the first year of follow up‡‡:								
No of deaths	608	129	111	72	71	63	63	1117
Mortality‡	1010	875	915	1010	1113	1327	1772	1072
Relative risk§	1.00	0.82*	0.83	0.83	0.87	1.03	1.29	—

*Two sided P< 0.05, **P<0.01, test for relative risk=1.0.
†The daily amounts (g) of ethanol corresponding to the six categories were ≤12.6, 12.7-24.9, 25.0-37.7, 37.8- 50.3, 50.4- 75.4, and ≥75.5.
‡Rate/100 000 person-years adjusted for age according to the person-year distribution of the entire cohort (≤ 54, 55-59, 60-64, and ≥65 years).
§Adjusted for age (≤54, 55-59, 60-64, and ≥65 years), level of education (primary school or below, secondary school, and college or higher), and cigarette smoking (non-smokers, ever smokers of ≤19 cigarettes/day, and ever smokers of ≥20 cigarettes/day).
¶Further adjustment for dietary factors.
††History of diabetes, emphysema, or hepatic cirrhosis.
‡‡All deaths and person-year contribution from the cohort during the 12 months after enrollment were excluded.

tively) and level of education (negatively) (data not shown).

As of 28 February 1995, the cohort had been followed up for 122 300 person-years (an average of 6.7 years of follow up per man). During the follow up 1198 of the men had died (980/100 000 person-years). There were 498 deaths from cancer, 269 deaths from stroke, and 104 deaths from ischaemic heart disease. We attempted to review the medical records of all cancer cases and deaths. Sixty per cent (298) of the deaths from cancer were confirmed histopathologically, including 93% (27) of deaths from colorectal cancer, 88% (80) from stomach cancer, 81% (26) from upper aerodigestive tract cancer, 67% (98) from lung cancer, and 13% (13) from liver cancer; 97% of the remaining cancer cases showed raised α fetoprotein concentrations or a positive liver scan, or both, at diagnosis.

Of the men who died from stroke, 85% (229) had been admitted to hospital before dying; no evidence of medical care immediately before death was established for 9% (24) of these patients. The comparable figures, respectively, were 82% (85) and 13% (13) for ischaemic heart disease, 84% (59) and 11% (8) for chronic bronchitis and emphysema, and 94% (33) and 3% (1) for hepatic cirrhosis.

Total mortality by alcohol consumption

Moderate drinkers had a reduced risk of death compared with non-drinkers or heavy drinkers. After age, level of education, and cigarette smoking were adjusted for, the relative risk of death in those drinking 14 or fewer drinks a week compared with lifelong non-drinkers was 0.81 (95% confidence interval 0.70 to 0.94). On the other hand, men who consumed 43 or more drinks a week (that is, six or more drinks a day) had a 30% excess risk of death (1.30; 1.01 to 1.68) compared with lifelong non-drinkers. We tested for a curvilinear (U shaped) relation between alcohol intake and total mortality, and the result was significant (P for quadratic effect $=0.001$). Further adjustment for dietary factors had minimal effect on the relative risks. When men who had given up drinking or who had potentially life threatening illnesses (diabetes, emphysema, or hepatic cirrhosis) were excluded the relation between alcohol intake and overall mortality did not change (table 2).

Previously, we had found that there were fewer deaths in the first year of follow up than in subsequent years.[18] In other words, terminally ill patients were less likely to participate in the cohort study than other eligible subjects. We repeated the analyses after excluding all deaths and person-years of follow up for the 12 months after recruitment. The relation between alcohol consumption and risk of death remained unchanged (table 2). We also examined the effect of length of follow up on the association (≤ 4 years v >4 years); no difference was detected (data not shown). Because excluding various groups of subjects did not affect our results all subsequent analyses were conducted on the entire cohort.

Table 3 presents the risk of mortality from all causes according to alcohol consumption and cigarette smoking. Within each smoking category, light to moderate drinkers (1-28 drinks a week) experienced a lower risk of death than lifelong non-drinkers or heavy drinkers (≥ 29 drinks a week). On the other hand, risk of death increased with increasing number of cigarettes smoked a day within each category of alcohol intake. The highest mortality risk was observed among heavy drinkers who also smoked cigarettes regularly.

Table 3 Relative risks* and 95% confidence intervals for total mortality by alcohol and cigarette consumption status at recruitment

	Non-drinkers	No of drinks/week of ever drinkers		
		1-14	15-28	≥29
Non-smokers:				
No of deaths	290	63	25	11
Relative risk (95% confidence interval)	1.00†	0.78 (0.59 to 1.02)	0.85 (0.56 to 1.28)	1.05 (0.57 to 1.92)
Ever smokers of<20 cigarettes/day:				
No of deaths	162	91	42	41
Relative risk (95% confidence interval)	1.21 (1.00 to 1.47)	0.94 (0.74 to 1.19)	0.93 (0.67 to 1.28)	1.79 (1.28 to 2.48)
Ever smokers of ≥20 cigarettes/day:				
No of deaths	204	100	85	84
Relative risk (95% confidence interval)	1.53 (1.28 to 1.83)	1.33 (1.06 to 1.67)	1.36 (1.06 to 1.74)	1.61 (1.26 to 2.07)

*Relative risks were adjusted for age (≤54, 55-59, 60-64, and ≥65 years) and level of education (primary school or below, secondary school, and college or higher).
†Reference group.

Table 4 Mortality from cancer by alcohol consumption status at recruitment

Cancer site (ICD-9)	Non-drinkers	No of drinks/week of ever drinkers 1-28	≥29	Total subjects
All sites:				
No of deaths	261	179	58	498
Mortality*	386	395	619	407
Relative risk†	1.00	0.85	1.08	—
Upper aerodigestive tract (141, 143-146, 148-150, 161):				
No of deaths	10	13	9	32
Mortality*	15	29	94	26
Relative risk†	1.00	1.49	3.72‡	—
Stomach (151):				
No of deaths	48	33	10	91
Mortality*	72	73	104	74
Relative risk†	1.00	0.98	1.37	—
Colon and rectum (153, 154):				
No of deaths	17	9	3	29
Mortality*	25	20	32	24
Relative risk†	1.00	0.80	1.30	—
Liver (155):				
No of deaths	61	32	9	102
Mortality*	90	70	104	83
Relative risk†	1.00	0.68	0.84	—
Lung (162):				
No of deaths	72	54	21	147
Mortality*	107	119	222	120
Relative risk†	1.00	0.76	0.94	—
All other sites:				
No of deaths	53	38	6	97
Mortality*	78	84	63	79
Relative risk†	1.00	0.98	0.66	—

*Rate/100 000 person-years adjusted for age according to the person-year distribution of the entire cohort (≤54, 55-59, 60-64, and ≥65 years).
†Relative risk, adjusted for age (≤54, 55-59, 60-64, and ≥65 years), level of education (primary school or below, secondary school, and college or higher), and cigarette smoking (non-smokers, ever smokers of ≤ 19 cigarettes/ day, and ever smokers of ≥20 cigarettes/ day).
‡Two sided P<0.01, test for relative risk=1.0.

Cause-specific mortality by alcohol consumption

Table 4 shows the relation between alcohol consumption and risk of death from all cancer and for the five commonest cancers. Cancer accounted for 42% of all deaths. After adjustment for age, level of education, and cigarette smoking, light to moderate drinkers had a non-significant 15% reduction in risk of death from any cancer relative to lifelong non-drinkers. As expected, heavy drinkers had a 3.7-fold increased risk of cancer of the upper aerodigestive tract, mainly oesophageal cancer (22). A non-significant 30-40% increase in risks of death from cancers of the stomach, colon, and rectum was observed in heavy drinkers.

Table 5 presents the association between alcohol intake and mortality from all non-cancer related causes and from the five commonest specific, non-cancer related causes. Stroke was the leading cause of non-cancer deaths in the study population, accounting for 22% of total deaths, while ischaemic heart disease accounted for only 9%. Compared with lifelong non-drinkers, light to moderate drinkers had a significant 20% reduction in mortality from all non-cancer related causes (0.80; 0.67 to 0.94), a 36% reduction in mortality from ischaemic heart disease (0.64; 0.41 to 0.998), and a 33% reduction in mortality from non-cancer related causes other than those specifically listed in table 5 (mainly other forms of heart disease and diabetes). Light to moderate drinking offered no protection from risk of death from stroke, and heavy drinking was associated with a significant 1.7-fold excess in risk of death from stroke (1.12 to 2.44). Adjustment for history of hypertension slightly diminished the effect of heavy drinking on risk of fatal stroke (1.53; 1.04 to 2.25). Heavy drinking was also positively associated with death from hepatic cirrhosis (2.99; 1.12 to 7.94).

We repeated all cause-specific analyses after excluding former drinkers, men with pre-existing serious illnesses, or deaths and person-year contribution from the cohort during the first 12 months after enrollment. The exclusions did not alter any of the relations described above. We also repeated all cause-specific analyses for the five commonest cancer sites after exclusion of cancer deaths without histopathological confirmation and for the five commonest non-cancer related causes after exclusion of deaths lacking evidence of medical care immediately before death. No material changes in results were observed.

Total mortality by type of alcoholic drink

Of the 7773 regular drinkers, 3500 (45%) drank beer (only 18 men reported consuming 29 or more drinks of beer a week), 4341 (56%) drank wine, and 3723 (48%) drank spirits. When we examined the relation between a specific type of alcoholic drink and overall mortality while simultaneously controlling for the total ethanol intake from the other two types as well as age,

Table 5 Mortality from non-cancer related causes by alcohol consumption status at recruitment

Non-cancer-related cause (ICD-9)	Non-drinkers	No of drinks/week of ever drinkers 1-28	≥29	Total subjects
All non-cancer-related causes:				
No of deaths	395	227	78	700
Mortality†	584	504	845	572
Relative risk‡	1.00	0.80**	1.20	—
Ischaemic heart disease (410-414):				
No of deaths	63	31	10	104
Mortality†	93	68	106	85
Relative risk‡	1.00	0.64*	0.88	—
Stroke (430-438):				
No of deaths	134	99	36	269
Mortality†	199	219	388	220
Relative risk‡	1.00	1.02	1.65*	—
Chronic bronchitis and emphysema (490-496):				
No of deaths	39	26	5	70
Mortality†	58	57	51	57
Relative risk‡	1.00	0.87	0.64	—
Hepatic cirrhosis (571):				
No of deaths	23	6	6	35
Mortality†	34	14	65	29
Relative risk‡	1.00	0.46	2.99*	—
Injury/accident (800-999):				
No of deaths	24	11	7	42
Mortality†	35	25	81	34
Relative risk‡	1.00	0.65	1.87	—
All other non-cancer related causes:				
No of deaths	112	54	14	180
Mortality†	165	121	154	147
Relative risk‡	1.00	0.67*	0.77	—

†Rate/100 000 person-years adjusted for age according to the person-year distribution of the entire cohort (≤54, 55-59, 60-64, and ≥65 years).
‡Adjusted for age (≤54, 55-59, 60-64, and ≥65 years), level of education (primary school or below, secondary school, and college or higher), and cigarette smoking (non-smokers, ever smokers of ≤ 19 cigarettes/ day, and ever smokers of ≥20 cigarettes/ day).
*Two sided P<0.05, **P<0.01, test for relative risk=1.0.

level of education, and smoking (model one, see methods for details), moderate drinkers, irrespective of the type of alcoholic drinks consumed, had a reduced risk of death relative to lifelong non-drinkers (table 6). Similarly, heavy drinkers of either wine or spirits showed a roughly 20% increase in mortality risk relative to non-drinkers. When we used model two (see methods) to test for possible type-specific effects on total mortality after accounting for total ethanol intake, the results were consistent with those from model one. We noted no additional effect from beer, wine, or spirits after total ethanol intake has been accounted for.

Discussion

We found that Chinese men who consumed no more than 14 drinks a week had a significant 19% reduction in overall mortality relative to lifelong non-drinkers. The observed association was not explicable by confounding factors such as cigarette smoking, dietary intake, level of education, and age. In agreement with findings in Western populations,[1-5 6 16 17] we observed a significant 36% reduction in mortality from ischaemic heart disease among light to moderate drinkers relative to lifelong non-drinkers. However, the protective effect of moderate drinking on total mortality in Shanghai men was not restricted to ischaemic heart disease, which accounts for only 9% of total deaths.

Bias is unlikely to explain the observed associations. The prospective study design precluded recall bias. A structured questionnaire was used to administer the interviews to ensure that all subjects were asked identical questions. We were also able to separate former drinkers from lifelong non-drinkers and thus eliminated the potential bias that former drinkers might have an increased risk of death and account for the raised mortality.[24] Another proposed argument to explain the relatively high mortality among abstainers is that such individuals carry a greater "burden of ill health" due to lifelong health problems than light to moderate drinkers, regardless of their previous drinking status.[25] We excluded all subjects with a history of potentially life threatening illnesses and no change in the results was observed. Our results also were unaffected by duration of follow up, supporting the notion that other unidentified bias is not likely to be responsible for the observed association.

Table 6 Total mortality by consumption status on specific types of alcoholic beverages at recruitment

		No of drinks/week of ever drinkers			
Type of drink	**Non-drinkers**	**≤7**	**8-14**	**15-28**	**≥29**
Beer:					
No of deaths	656	146	41	8	0
Mortality*	971	875	830	1284	—
Relative risk†	1.00	0.88	0.78	0.96	—
Wine:					
No of deaths	656	118	75	73	36
Mortality*	971	784	1058	1082	1586
Relative risk†	1.00	0.73	0.99	0.82	1.22
Spirits:					
No of deaths	656	60	70	76	83
Mortality*	971	774	1106	1016	1628
Relative risk†	1.00	0.71‡	0.90	0.86	1.17

*Rate/100 000 person-years adjusted for age according to the person-year distribution of the entire cohort (≤54, 55-59, 60-64, and ≥65 years).
†Relative risk, adjusted for age (≤54, 55-59, 60-64, and ≥65 years), level of education (primary school or below, secondary school, and college or higher), and cigarette smoking (non-smokers, ever smokers of ≤19 cigarettes/ day, and ever smokers of ≥20 cigarettes/ day), and consumption of other two types of alcoholic beverages.
‡Two sided P<0.01, test for relative risk=1.0.

Key messages

- Moderate alcohol consumption is associated with a reduced risk of death from all causes and ischaemic heart disease in Western populations
- No data are available for Chinese people, who drink little grape wine and have a low rate of ischaemic heart disease
- In this study of middle aged Chinese men regular drinkers of small amounts of alcohol had a 19% lower death rate than non-drinkers
- Death rates among moderate drinkers were lower for cancer and non-cancer causes, and the type of alcohol made no difference
- Relative to non-drinkers, heavy drinkers had a 30% increased risk of death

There is evidence that the collected information is reasonably reliable. Diseases reported to be positively associated with heavy alcohol intake in Western populations (upper aerodigestive cancer, liver cirrhosis, and injury/accident) showed similar associations in our study. It is also important to note that loss to follow up was negligible (only 108 out of 18 244 subjects). Finally, it is unlikely that the observed associations between cause-specific mortality and alcohol intake were due to misclassification of cause of death. In 82% of cases cause of death was medically confirmed immediately before death, and exclusion of those without such confirmation did not affect the association between light to moderate alcohol intake and cause-specific mortality.

An ecological study suggested that high consumption of wine might be responsible for the lower risk of ischaemic heart disease in France.[12] Recent findings from a prospective study in Copenhagen have shown that wine drinkers had a 50% reduction in mortality relative to non-drinkers.[13] In the same study, beer drinking was not related to mortality and spirit consumption increased the risk.[13] However, several studies have reported no differential effects of wine relative to other types of alcoholic beverages on mortality.[11 14-17] Rimm *et al* recently concluded that much of the benefit is from alcohol rather than other components of each type of drink.[18] We did not separate the grape wine from other wines drunk by Chinese, but data from a case-control study of nasopharyngeal carcinoma that we conducted in Shanghai in the late 1980s indicate that few Chinese men drink grape wine regularly. Among male controls in that study, only 6% of wine drinkers drank grape wine; the rest of them drank only rice wine (unpublished data). Our data did not show that intake of wine was more beneficial in reducing risk of death than consumption of beer or spirits.

In summary, light drinking was associated with a 19% reduction in overall mortality in middle aged men in Shanghai. This protective effect was not restricted to any specific type of alcoholic drink. There was also a

36% reduction in risk of death from ischaemic heart disease among light to moderate drinkers and significant increases in risks of cancer of the upper aerodigestive tract, stroke, and hepatic cirrhosis among heavy drinkers.

We thank X-L Wang, Y-L Zhang, and J-R Cheng of the Shanghai Cancer Institute for their help in data collection and management, the staff at the Shanghai Cancer Registry for their help in verifying cancer diagnosis, and Kazuko Arakawa of the University of Southern California for her help in data management and analysis.

Funding: Grants R01 CA43092 and R35 CA53890 from the National Cancer Institute, Bethesda, Maryland, USA.

Conflict of interest: None.

1 Doll R, Peto R, Hall E, Wheatley K, Gray R. Mortality in relation to consumption of alcohol: 13 years' observations on male British doctors. *BMJ* 1994;309:911-8.
2 Gronbaek M, Deis A, Sorensen TIA, Becker U, Borch-Johnsen K, Muller C, *et al.* Influence of sex, age, body mass index, and smoking on alcohol intake and mortality. *BMJ* 1994;308:302-6.
3 Boffetta P, Garfinkel L. Alcohol drinking and mortality among men enrolled in an American Cancer Society Prospective Study. *Epidemiology* 1990;1:342-8.
4 Rehm J, Sempos CT. Alcohol consumption and all-cause mortality. *Addiction* 1995;90:471-80.
5 Serdula MK, Koong S-L, Williamson DF, Anda RF, Madans JH, Kleinman JC, Byers T. Alcohol intake and subsequent mortality: Findings from the NHANES I follow-up study. *J Stud Alcohol* 1995;56:233-9.
6 Cullen KJ, Knuiman MW, Ward NJ. Alcohol and mortality in Busselton, Western Australia. *Am J Epidemiol* 1993;137:242-8.
7 Goldberg RJ, Burchfiel CM, Reed DM, Wergowske G, Chiu D. A prospective study of the health effects of alcohol consumption in middle-aged and elderly men: the Honolulu Heart Program. *Circulation* 1994;89:651-9.
8 Farchi G, Fidanza F, Mariotti S, Menotti A. Alcohol and mortality in the Italian rural cohorts of the seven countries study. *Int J Epidemiol* 1992;21:74-81.
9 De Labry LO, Glynn RJ, Levenson MR, Hermos JA, LoCastro JS, Vokonas PS. Alcohol consumption and mortality in an American male population: recovering the U-shaped curve-findings from the normative aging study. *J Stud Alcohol* 1992;53:25-32.
10 Fuchs CS, Stampfer MJ, Colditz GA, Giovannucci EL, Manson JE, Kawachi I, *et al.* Alcohol consumption and mortality among women. *N Engl J Med* 1995;332:1245-50.
11 Friedman LA, Kimball AW. Coronary heart disease mortality and alcohol consumption in Framingham. *Am J Epidemiol* 1986;124:481-9.
12 Renaud S, De Lorgeril M. Wine, alcohol, platelets, and the French paradox for coronary heart disease. *Lancet* 1992;339:1523-6.
13 Gronbaek M, Deis A, Sorensen TIA, Becker U, Schnohr P, Jesen G. Mortality associated with moderate intakes of wine, beer, and spirits. *BMJ* 1995;310:1165-9.
14 Yano K, Rhoads GG, Kagan A. Coffee, alcohol and risk of coronary heart disease among Japanese men living in Hawaii. *N Engl J Med* 1977;297:405-9.
15 Stampfer MJ, Colditz GA, Willett WC, Speizer FE, Hennekens CH. A prospective study of moderate alcohol consumption and the risk of coronary disease and stroke in women. *N Engl J Med* 1988;319:267-73.
16 Klatsky AL, Armstrong MA, Friedman GD. Alcohol and mortality. *Ann Intern Med* 1992;117:646-54.
17 Rimm EB, Giovannucci EL, Willett WC, Colditz GA, Ascherio A, Rosner B, Stampfer MJ. Prospective study of alcohol consumption and risk of coronary disease in men. *Lancet* 1991;338:464-8.
18 Rimm EB, Klatsky A, Grobbee D, Stampfer MJ. Review of moderate alcohol consumption and reduced risk of coronary heart disease: is the effect due to beer, wine, or spirits? *BMJ* 1996;312:731-6.
19 Yuan J-M, Ross RK, Wang XL, Gao YT, Henderson BE, Yu MC. Morbidity and mortality in relation to cigarette smoking: a prospective male cohort study in Shanghai, China. *JAMA* 1996;275:1646-50.
20 Adames CF. Nutritive value of American foods in common unit. *Agriculture Handbook No 456.* Washington, DC: US Department of Agriculture, US Government Printing Office, 1975:31.
21 Ross RK, Yuan J-M, Yu MC, Wogan GN, Qian G-S, Tu J-T, Groopman JD, *et al.* Urinary aflatoxin biomarkers and risk of hepatocellular carcinoma. *Lancet* 1992;339:943-6.
22 Breslow NE, Day NE. *Statistical Method in Cancer Research, Volume II: The design and analysis of cohort studies.* Lyon, France: International Agency for Research on Cancer, 1987:120-76. (IARC scientific publications No 82.)
23 Cox DR. Regression models and life tables. *Journal of the Royal Statistical Society* 1972;34(B):187-220.
24 Shaper AG, Wannamethee G, Walker M. Alcohol and mortality in British men: explaining the U-shaped curve. *Lancet* 1988;ii:1267-73.
25 Shaper AG. Alcohol and mortality: a review of prospective studies. *Br J Addict* 1990;85:837-47.

(Accepted 16 October 1996)

Underestimation and undertreatment of pain in HIV disease: multicentre study

François Larue, Alain Fontaine, Sophie M Colleau

Abstract

Objectives: To measure the prevalence, severity, and impact of pain on quality of life for HIV patients; to identify factors associated with undertreatment of pain.

Design: Multicentre cross sectional survey.

Settings: 34 HIV treatment facilities, including inpatient hospital wards, day hospitals, and ambulatory care clinics, in 13 cities throughout France.

Subjects: 315 HIV patients at different stages of the disease.

Main outcome measures : Patients: recorded presence and severity of pain and rated quality of life. Doctors: reported disease status, estimate of pain severity, and analgesic treatment ordered.

Results: From 30% (17/56) of outpatients to 62% (73/118) of inpatients reported pain due to HIV disease. Pain severity significantly decreased patients' quality of life. Doctors underestimated pain severity in 52% (70/135) of HIV patients reporting pain. Underestimation of pain severity was more likely for patients who reported moderate (odds ratio 24) or severe pain (165) and less likely for patients whose pain source was identified or who were perceived as more depressed. Of the patients reporting moderate or severe pain, 57% (61/107) did not receive any analgesic treatment; only 22% (23/107) received at least weak opioids. Likelihood of analgesic prescription increased when doctors estimated pain to be more severe and regarded patients as sicker.

Conclusions: Pain is a common and debilitating symptom of HIV disease which is gravely underestimated and undertreated.

Consultation de Traitement de la Douleur, Institut Mutualiste Montsouris, 42 Boulevard Jourdan, 75014 Paris, France

François Larue, *pain consultant*

Département de Santé Publique, Faculté Bichat, Université Paris VII, 75018 Paris, France

Alain Fontaine, *health services research staff physician*

BMJ 1997;314:23–8

continued over

Introduction

In the past 15 years intensive clinical research has produced therapeutic improvements for patients infected with HIV, but issues of control of pain and symptoms in such patients have only recently begun to be studied.

WHO Collaborating Center for Symptom Evaluation in Cancer Care, University of Wisconsin Medical School, Madison, WI 53705, USA

Sophie M Colleau, *assistant director for communication*

Correspondence to: Dr Larue.

Estimates of prevalence of pain during HIV infection range between 30% and 80%.[1-6] Data, however, have mostly been collected in single site studies,[1 3 5] for patients with full blown AIDS,[3-6] or for those in the terminal stage of the disease,[4] and are sometimes based on review of patients' records.[3 6]

Pain experienced by HIV patients can be due to multiple sources[7]: firstly, the HIV infection itself or its consequences (infections, tumours); secondly, treatments for AIDS; or, thirdly, it can be unrelated to the disease and its treatment. HIV related pain has been found to impair both functional and affective components of daily life.[1 2 8 9]

It has been recommended that pain management for AIDS patients follows the model of the World Health Organisation analgesic ladder developed for the management of cancer pain.[7 10-12] Although this model has not yet been systematically validated, clinical reports describe its successful application to the management of pain in AIDS, particularly with respect to the use of opioids to treat severe pain.[13 14] In addition, adjuvant drugs, such as antidepressants, have been shown to be effective for the treatment of neuropathic pain,[10] which seems common in the course of AIDS,[7] and is only partly sensitive to opioids.[15]

A few studies have suggested that pain management for HIV patients is inadequate.[1 3 16] A recent study of ambulatory patients with AIDS in New York City showed that characteristics of the patients themselves (sex, education, use of injected drugs) can be associated with higher risk of undertreatment.[16] Factors responsible for the undertreatment of cancer pain include doctors' underestimation of patients' pain and lack of knowledge regarding proper analgesic treatment.[12 17 18] The role of these factors in the management of pain in HIV has not been assessed.

We designed this multicentre study to determine the proportion of HIV patients who experience significant pain, to assess the impact of pain severity on quality of life, to describe pain treatment practices, and to identify factors associated with undertreatment.

Patients and methods

Between 20 September 1992 and 10 October 1992 we studied 315 HIV patients drawn from 34 treatment settings in 13 cities throughout France. Participating institutions were 13 inpatient hospital wards, 12 day hospitals, and nine ambulatory care clinics. All were Centres for Information and Treatment of Human Immunodeficiency (CISIH), which are the basic administrative and research units for the management of HIV disease in France. Institutional participation in the study was voluntary and depended on the availability of local coordinators trained by the AIDS Task Force of the French Ministry of Health, which served as liaison among study sites. All received specific instruction regarding the study protocol.

Patients and doctors

All HIV patients aged 18 or over, irrespective of stage of disease, were eligible, except those unable to fill out the questionnaire because they either did not speak French or presented cognitive impairment.

All eligible inpatients admitted during the study period were surveyed once within the first three days after admission. In day hospitals or ambulatory care clinics one patient was included every day: if n patients were expected in either treatment setting on a given day, the nth/2 patient was asked to participate; if a patient refused or was ineligible the next patient was considered; when patients were available at both day hospital and ambulatory care clinic in the same city the study was conducted in each setting on alternative days.

Patients completed the French language version of the brief pain inventory,[19] previously validated in a study of pain among cancer patients.[20 21] The inventory has also been used previously to measure pain in patients with AIDS.[1 9 16] It asks patients to report if they have experienced pain because of their disease during the previous week and to rate their pain (at its worst, at its least, and on average) on 0-10 numerical scales. In addition, patients were asked to rate their quality of life (0-10 numerical rating scale) and a set of symptoms including fatigue, sadness, and depression (4 point verbal rating scales).

Each doctor was asked to describe the main characteristics of his or her patient's disease and to assess the patient's Karnofsky performance status. Doctors were also asked to rate their patients' severity of pain and level of anxiety and depression, to identify the source of pain if known, and to describe the patient's current analgesic treatment.

All questionnaires were distributed and collected at each site by the local coordinator. Patients were asked to sign an informed consent form after reading the study's objectives. Patients answered the questionnaires without any medical supervision but could be assisted by the local coordinator. Outpatients were asked to complete the questionnaire before their consultation. Doctors were asked to complete the questionnaire within two days of seeing the patient, without knowing the patient's responses.

Patients were classified in three groups (asymptomatic, A1-A2; pathological, B1-B2; full blown AIDS, A3, B3, C1-3) according to the Centers for Disease Control.[22]

A patient was considered to have pain if he or she reported experiencing pain due to his or her disease in the past week or rated pain severity above 0 on the worst pain scale, or both. Worst pain rated 5 or above was defined as significant pain, because such pain has been shown to interfere significantly with function.[21 23]

Analysis of treatment

Following WHO guidelines[12] we examined analgesic treatments with respect to pain severity reported by the patients. We categorised the patient's worst pain score as mild (1-4), moderate (5-6), and severe (7-10),[21] and we constructed a pain management index as in previous studies.[20 23] Negative scores on the index are considered a conservative indicator of undertreatment.

Actually, undertreatment of pain could reflect underestimation of pain or underprescription of analgesics, or both. Because patients' reports of average pain were the most highly correlated with doctors' estimates, we scored the patients' reports of average pain on a 3 point scale (mild, moderate, severe) according to the same cutpoints as those used to categorise worst pain. The doctors' estimate of the patients' pain was scored on a 4 point scale (no pain, mild, moderate,

Table 1 Patients' characteristics by treatment setting. Figures are numbers (percentages) of patients

Characteristic	In-patients (n=118)	Day patients (n=116)	Out-patients (n=56)	P value	Others* (n=25)
Infection:				0.39	
Homosexuality	56(48)	53(46)	28(50)		5
Injecting drug misuse	33(28)	29(25)	12(21)		11
Homosexuality and injecting drug misuse			1 (2)		
Heterosexuality	15(13)	25(22)	11(20)		1
Haemophilia or blood infusion	7 (6)	8 (7)	2 (4)		
Unknown	7 (6)	1 (1)	2 (4)		8
Centers for Disease Control classification:				< 0.00001	
Asymptomatic	4 (3)	4 (3)	11(20)		1
Pathological	6 (5)	19(16)	13(23)		3
Full blown AIDS	95(80)	87(75)	28(50)		14
Missing	13(11)	6 (5)	4 (7)		7
Karnofsky performance status:				< 0.00001	
≥ 80%	34(29)	59(51)	43(77)		11
40-70%	66(56)	55(47)	13(23)		9
≤ 30%	12(10)	2 (2)			
Missing	6 (5)				5
Pain in past week	73(62)	61(53)	17(30)	0.0005	13
Significant pain†	60(51)	38(33)	9(16)	0.00006	11
Ratio of significant pain to pain in past week	82%	62%	53%	0.007	85%

*Treatment setting not reported.
†Patients report worst pain ≥5.

severe). Underestimation of pain severity was defined as doctors rating pain lower than their patients on these ordinal scales. To measure the congruence between the doctors' rating of the patients' pain and their prescribing behaviour, we then constructed a prescription adequacy index, calculated like the pain management index but based on the doctors' rating of the patients' pain rather than on the patients' self reports. Negative values on the prescription adequacy index indicate that the medications prescribed do not match the level of pain assessed by the doctor.

Descriptive statistics were produced to examine patients' characteristics, pain prevalence, and pain severity in each setting. The impact of pain on patients' self reported quality of life was assessed in a multiple linear regression model. Control variables, including treatment setting, stage of disease, fatigue, sadness, and depression, were selected for face validity. Logistic regression analyses identified factors associated with underestimation of pain severity for patients reporting pain and factors associated with pain treatment. Dependent variables were doctors' underestimation of pain and doctors' prescription of any analgesic medication, regardless of its potency. Independent variables included the patients' own reports of pain and the doctors' ratings of the patients' pain severity, as well as patients' characteristics and treatment setting. For each regression, a reduced model was produced by backward elimination. Significance was assessed from the reduction in the goodness of fit of the model.[24] Results are expressed in terms of the odds ratio which estimates how each independent variable affects the probability of the behaviour described by the dependent variable.

Results

A total of 118 questionnaires were collected from inpatient hospital wards, 116 from day hospitals, and 56 from ambulatory care clinics. Only two patients refused to participate in the study. The treatment setting was not reported for 25 patients, whose characteristics were similar to those of the other patients; these patients were excluded from the analysis.

Patients' ages ranged from 21 to 66 years (median 33); the sample was essentially male (78%). Table 1 shows characteristics related to disease and pain. Patients did not differ significantly between settings with respect to the source of infection. Inpatients had more advanced HIV infection than day patients or outpatients.

Prevalence, severity, and source of pain—Pain was significantly more common for inpatients (62%; 73/118) than for day patients (53%; 61/116) and outpatients (30%; 17/56) (P = 0.0005). The prevalence of significant pain (worst pain level ≥ 5) ranged from 16% (9/56) for outpatients to 51% (60/118) for inpatients. Doctors were able to identify at least one source of pain for over two thirds of the patients reporting pain (105/151). Among 78 AIDS patients for whom a source of pain was identified, 26 (33%) were reported to have digestive or mouth pain, 25 (32%) muscular pain, and 16 (20%) joint or bone pain. Central nervous system pain was reported for 15 patients (19%) and painful peripheral neuropathies for 10 (13%).

Impact of pain on quality of life—Patients with significant pain (worst pain level ≥ 5) reported lower quality of life during the week before the survey than patients with no pain (fig 1). The linear regression of worst pain on quality of life scores confirms that significant pain had an independent negative impact on HIV patients' quality of life, after adjustment for treatment setting, stage of the disease, fatigue, sadness, and depression (table 2).

Adequacy of pain estimation—Both doctors' and patients' pain ratings were available for 135 patients reporting an average pain level greater than zero. Doctors underestimated pain severity in 52% (70) of these patients. Table 3 presents predictors of underestimation. Underestimation was much more likely for

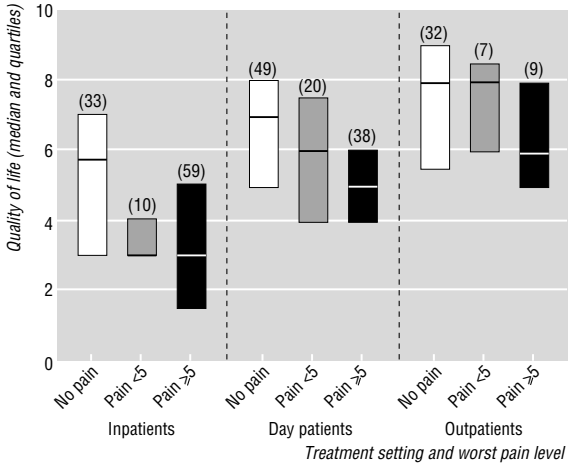

Fig 1 Quality of life and pain by treatment setting. Numbers are numbers of patients in each group

Table 2 Impact of HIV related pain on quality of life in 219 patients

Detail	Coefficient (SE)	P value
Full blown AIDS (0-1)	–0.32 (0.60)	0.59
Pathological, B1B2 (0-1)	–0.23 (0.69)	0.74
Day patient (0-1)	–0.92 (0.42)	0.03
Inpatient (0-1)	–1.65 (0.46)	0.0004
Patient's rating of fatigue (0-3)	–0.57 (0.19)	0.003
Patient's rating of sadness (0-3)	–0.37 (0.21)	0.07
Patient's rating of depression (0-3)	–0.57 (0.22)	0.01
Patient reports worst pain<5 (0-1)	–0.56 (0.44)	0.20
Patient reports worst pain ≥5 (0-1)	–0.73 (0.35)	0.04
Constant	8.83 (0.61)	< 0.0001

r^2=0.39 (Multiple linear regression with adjustment for stage of disease, treatment setting, fatigue, sadness, and depression).

patients with severe (odds ratio 165) or moderate pain (24) than for patients with mild pain and somewhat less likely when the patient was perceived as more depressed or when doctors could identify the source of pain.

Adequacy of pain treatment—Table 4 shows the types of analgesic medication prescribed according to the patients' worst pain reports. While 69 patients reported severe pain, only 10 (15%) received opioids and almost half did not receive any analgesic treatment. Among 38 patients reporting moderate pain, only 5 (13%) received at least weak opioids, and most did not receive any analgesic treatment. Overall, 85% (123/144) of the patients whose worst pain level was greater than zero were undermedicated (had negative pain management index scores) according to the WHO guidelines. When severity of pain estimated by the doctor was used as the reference, pain treatment remained inadequate (negative prescription adequacy index scores) for 70% (85/121) of the patients. Antidepressants are considered useful for the treatment of neuropathic pain.[10] In this study antidepressants were prescribed to 21 patients but to only one of the 33 patients suffering painful peripheral neuropathies and to two of the 29 patients with central nervous system pain. Table 5 identifies factors associated with the prescription of any analgesic medication, regardless of its potency. Likelihood of prescription increased as doctors recognised pain and estimated it to be more severe; injecting drug users, homosexuals, and patients who seemed more sick (Karnofsky ≤ 70%) were also more likely to receive some type of analgesic. Identification of the source of pain by the doctor was significantly associated with analgesic treatment in bivariate analyses.

Discussion

This survey is the first multicentre study of pain in HIV patients at all stages of the disease. It was conducted in France, the country with the highest number of HIV patients in Europe. Our sample is not strictly representative of the general population of such patients in France as participation of institutions was voluntary and patients at advanced stages of the disease were over-represented.[25]

Pain prevalence and severity were high, comparable with results from a French nationwide study of pain in patients with cancer.[20] The higher prevalence and severity of pain observed among HIV inpatients suggest that pain increases as the disease progresses, as shown in a previous prospective study.[2] While we recognise limitations in our measure of quality of life, our data show, most importantly, that significant pain

Table 3 Predictors of underestimation of pain by doctors in 116 patients with pain

		Initial logistic regression model			Reduced logistic regression model		
Predictor	Bivariate P	Regression coefficient (SE)	Adjusted odds ratio (95% confidence interval)	P value*	Regression coefficient(SE)	Adjusted odds ratio (95% confidence interval)	P value*
Patient reports severity of average pain (reference category: weak):	< 0.00001†			< 0.0001			< 0.0001
As moderate		3.386 (0.814)	29.6 (6.0 to 145.7)	< 0.0001	3.167 (0.704)	23.7 (6.0 to 94.3)	< 0.0001
As severe		5.539 (1.129)	254.5 (27.8 to 2325.8)	< 0.0001	5.110 (0.978)	165.0 (24.4 to 1126.5)	< 0.0001
Doctor rates patient's depression (0-10)	0.10‡	–0.314 (0.139)	0.7 (0.6 to 1.0)	0.024	–0.282 (0.106)	0.7 (0.6 to 0.9)	0.008
Doctor identifies pain source (0-1)	0.19†	–1.432 (0.675)	0.2 (0.1 to 0.9)	0.034	–1.286 (0.622)	0.3 (0.1 to 0.9)	0.038
CD4 cells count (reference category: > 350):	0.74†			0.133			0.089
200-349		0.627 (1.143)	1.9 (0.2 to 17.6)	0.583	0.322 (1.030)	1.4 (0.2 to 10.4)	0.755
50-199		–0.006 (0.986)	1.0 (0.1 to 6.9)	0.995	–0.379 (0.892)	0.7 (0.1 to 3.9)	0.671
< 50		–1.472 (1.041)	0.2 (0.0 to 1.8)	0.157	–1.561 (0.881)	0.2 (0.0 to 1.2)	0.076
Doctor reports patient sought care because of pain (0-1)	0.92†	–0.951 (0.707)	0.4 (0.1 to 1.5)	0.179	–1.061 (0.667)	0.3 (0.1 to 1.3)	0.112
Setting (reference category: outpatients):	0.09†			0.488			
Day patients		–0.817 (0.928)	0.4 (0.1 to 2.7)	0.378			
Inpatients		–0.209 (0.987)	0.8 (0.1 to 5.6)	0.832			
Doctor rates patient's anxiety (0-10)	0.05‡	–0.015 (0.135)	1.0 (0.8 to 1.3)	0.913			
Contamination (reference category: others):	0.21†			0.555			
Homosexual		0.509 (0.768)	1.7 (0.4 to 7.5)	0.507			
Injecting drug misuser		0.87 (0.803)	2.4 (0.5 to 11.5)	0.279			
Age	0.46‡	–0.017 (0.035)	1.0 (0.9 to 1.1)	0.626			
Karnofsky (reference category: 80-100%)	0.22†			0.397			
40-70%		–0.73 (0.637)	0.5 (0.1 to 1.7)	0.251			
≤ 30%		0.311 (1.332)	1.4 (0.1 to 18.6)	0.816			
Sex	0.75†	1.037 (0.769)	2.8 (0.6 to 12.7)	0.178			
Constant		0.23 (1.684)		0.89	0.305 (0.892)		0.73

*Likelihood ratio statistic. †χ^2 Test. ‡Spearman correlation.
All variables were entered into logistic regression model.
Final model produced by backward elimination.

Table 4 Pain treatment by pain severity

Treatment and management indices	Worst pain reported by patients: Mild (n=37)	Moderate (n=38)	Severe (n=69)	Total (n=144)	P value
No analgesic treatment	31	29	32	92	
Non-opioid analgesic	3	4	19	26	
Weak opioid	3	4	8	15	
Morphine	0	1	10	11	
Negative pain management index (patients' pain rating)	31 (84%)	33 (87%)	59 (85%)	123 (85.4%)	0.93
Negative prescription adequacy index (doctors' rating of patients' pain)	24/30 (80%)	23/30 (77%)	38/61 (62%)	85/121 (70%)	0.15

diminishes the quality of life of patients with HIV disease, regardless of treatment setting and stage of disease.

This study also shows that HIV related pain was both severely underestimated and undermedicated. It is especially disturbing that doctors underestimated pain when it was most severe. Moreover, doctors seemed reluctant to prescribe potent medications, even though they took into account their rating of the patients' pain to prescribe analgesics. They also failed to use antidepressants when pain relief could be expected from those adjuvant drugs. Inappropriate pain treatment has previously been shown in France[20 26] and may be related to insufficient awareness and training regarding pain management.[27] Insufficient knowledge and reluctance to use morphine constitute barriers to adequate pain management in cancer.[18] The same factors may apply here. Interestingly, injected drug use did not seem to be an obstacle to the prescription of analgesics.

We found that the ability of a clinician to identify the source of pain led to better pain assessment and, possibly, to its treatment. The subjectivity of pain in a complex disease such as HIV infection where somatic, visceral, and neuropathic pains are often associated with anxiety and depression makes recognition and assessment of pain particularly daunting for clinicians.[2 7 8] Pain assessment in HIV patients may require development of specific methods, including multisymptom assessment scales. Recognising the validity of the patients' point of view still remains the cornerstone of adequate pain assessment.

Table 5 Predictors of analgesic treatment in 246 patients with HIV disease

Predictors	Bivariate P value	Initial logistic regression model: Regression coefficient (SE)	Adjusted odds ratio (95% confidence interval)	P value*	Reduced logistic regression model: Regression coefficient(SE)	Adjusted odds ratio (95% confidence interval)	P value*
Doctor rates patient's pain (reference category: no pain):	< 0.00001†			0.010			0.002
As weak		1.824 (0.718)	6.2 (1.5 to 25.3)	0.011	1.848 (0.698)	6.3 (1.6 to 24.9)	0.008
As moderate		2.218 (0.860)	9.2 (1.7 to 49.6)	0.010	2.345 (0.821)	10.4 (2.1 to 52.2)	0.004
As severe		3.074 (0.926)	21.6 (3.5 to 132.8)	0.001	3.162 (0.857)	23.6 (4.4 to 126.7)	0.0002
Contamination (reference category: others):	0.075†			0.009			0.012
Homosexual		0.955 (0.530)	2.6 (0.9 to 7.3)	0.072	0.913 (0.483)	2.5 (1.0 to 6.4)	0.059
Injecting drug misuser		1.742 (0.568)	5.7 (1.9 to 17.4)	0.002	1.574 (0.531)	4.8 (1.7 to 13.7)	0.003
Karnofsky (reference category: 80-100%):	< 0.001†			0.016			0.014
40-70%		0.868 (0.446)	2.4 (1.0 to 5.7)	0.052	0.872 (0.385)	2.4 (1.1 to 5.1)	0.023
≤ 30%		2.017 (0.840)	7.5 (1.4 to 39.0)	0.016	1.922 (0.784)	6.8 (1.5 to 31.8)	0.014
Doctor identifies pain source(s) (0-1)	< 0.00001†	0.579 (0.487)	1.8 (0.7 to 4.6)	0.235	0.668 (0.469)	1.9 (0.8 to 4.9)	0.154
Doctor rates patient's anxiety (0-10)	0.011‡	0.128 (0.091)	1.1 (1.0 to 1.4)	0.158			
Doctor rates patient's depression (0-10)	0.073‡	–0.090 (0.092)	0.9 (0.8 to 1.1)	0.326			
Doctor reports patient sought care because of pain (0-1)	< 0.001†	0.500 (0.510)	1.6 (0.6 to 4.5)	0.328			
CD4 cells count (reference category: > 350):	0.124†			0.986			
200-349		–0.176 (0.787)	0.8 (0.2 to 3.9)	0.824			
50-199		–0.175 (0.709)	0.8 (0.2 to 3.4)	0.805			
< 50		–0.258 (0.694)	0.8 (0.2 to 3.0)	0.710			
Setting (reference category: outpatients):	0.093†			0.925			
Inpatient		0.234 (0.612)	1.3 (0.4 to 4.2)	0.697			
Day patient		0.154 (0.594)	1.2 (0.4 to 3.7)	0.796			
Sex	0.680†	0.163 (0.512)	1.2 (0.4 to 3.2)	0.751			
Age	0.914‡	0.009 (0.022)	1.0 (1.0 to 1.1)	0.677			
Constant		–5.594 (1.335)		< 0.0001	–4.864 (0.772)		< 0.0001

*Likelihood ratio statistic.
†χ^2 Test.
‡Spearman correlation.
All variables entered into logistic regression model.
Final model produced by backward elimination.

Key messages

- Pain is a common and debilitating symptom of HIV disease; it is seriously undertreated
- This multicentre study shows that pain is present in 62% of HIV inpatients, that its severity decreases their quality of life, and that over half with significant pain do not receive any analgesic treatment
- Undertreatment of pain in HIV disease is related to doctors both underestimating pain and underprescribing analgesics
- The more severe the pain, the more often doctors underestimate it
- Doctors are reluctant to prescribe potent analgesics. Likelihood of analgesic prescription increases when doctors estimate pain to be more severe and regard patients as sicker

Improvement of management of pain for HIV patients will require new public health strategies to promote necessary changes in clinicians' knowledge, attitudes, and practices.

We thank L Bonifassi and Dr R Demeulemeester for their help in the study protocol, and Drs Ch Michon, Ph Vinceneux, J Pouchot, JM Nadal, and G Bez for their critical reading of the manuscript. We also thank the participating institutions: CISIH of Clamart, Colombes, Dijon, Grenoble, Marseille, Montpellier, Nice, Paris (Bichat), Rennes, Rouen, Strasbourg, Toulouse (Rangueil and Purpan), and Tourcoing; and L Brasseur, M Couturier, F Cornu, and M Mullier of Action Douleur for their help in pursuing initiatives to improve pain control in France.

Funding: Foundation de l'Avenir pour la Recherche Médicale Appliquée (grant No: 92/76) and the AIDS Task Force of the French Ministry of Health.

Conflict of interest: None.

1 McCormack JP, Li R, Zarowny D, Singer J. Inadequate treatment of pain in ambulatory HIV patients. *Clin J Pain* 1993;9:279-83.
2 Singer EJ, Zorilla C, Fahy-Chandon B, Chi S, Syndulko K, Tourtelotte WW. Painful symptoms reported by ambulatory HIV infected men in a longitudinal study. *Pain* 1993;54:15-9.
3 Lebovits AH, Smith G, Maignan M, Lefkowitz M. Pain in hospitalized patients with AIDS: analgesic and psychotropic medications. *Clin J Pain* 1994;10:156-61.
4 Shofferman J. Pain: diagnosis and management in the palliative care of AIDS. *J Palliat Care* 1988;4:46-9.
5 Moss V. Patient characteristics, presentation and problems encountered in advanced AIDS in a hospice setting: a review. *Palliat Med* 1991;5:112-6.
6 Dixon P, Higinson A. AIDS and cancer pain treated with slow release morphine. *Postgrad Med J* 1991;67:S92-4.
7 O'Neill WM, Sherrard JS. Pain in immunodeficiency virus disease: a review. *Pain* 1993;54:3-14.
8 Breitbart W. Suicide risk and pain in cancer and AIDS patients. In: Chapman CR, Foley KM, eds. *Current and emerging issues in cancer pain: research and practice*. New York: Raven Press, 1993:49-65.
9 Hoyt MJ, Nokes K, Newshan G, Staats JA, Thorn M. The effect of chemical dependency on pain perception in persons with AIDS. *J Assoc Nurses AIDS Care* 1994;5:33-8.
10 Jacox A, Carr DB, Payne R and the Management of Cancer Pain Guideline Panel. *Clinical practice guideline number 9: Management of cancer pain*. Rockville, Marylnad: US Department of Health and Human Services; Agency for Health Care Policy and Research, 1994.
11 Dormont J. *Prise en charge des personnes atteintes par le VIH. Rapport au Ministre*. Paris: Flammarion Medecine Science, 1993.
12 World Health Organisation. Cancer pain relief and palliative care. *WHO Tech Rep Ser* 1990; No 804.
13 Anand J, Carmosino L, Glatt AE. Evaluation of recalcitrant pain in HIV infected hospitalized patients. *J AIDS* 1994;7:52-6.
14 Newshan GT, Wainapel SF. Pain characteristics and their management in persons with AIDS. *J Assoc Nurses AIDS Care* 1993;4:53-9.
15 Portenoy RK, Foley KM, Inturisi CE. The nature of opioid responsiveness and its implications for neuropathic pain: new hypotheses derived from studies of opioid infusions. *Pain* 1990;**43**:273-86.
16 Breitbart W, Rosenfeld BD, Passik SD, McDonald MV, Thaler H, Portenoy RK. The undertreatment of pain in ambulatory AIDS patients. *Pain* 1996;65:243-9.
17 Von Roenn JH, Cleeland CS, Gonin R, Hatfield AK, Pandya KJ. Physician attitude and practice in cancer pain management: a survey from the Eastern Cooperative Oncology Group. *Ann Intern Med* 1993;119:121-6.
18 Larue F, Colleau SM, Fontaine A, Brasseur L. Oncologists and primary care physicians' attitudes toward pain control and morphine prescribing in France. *Cancer* 1995;76:2375-82.
19 Cleeland CS. Measurement of pain by subjective report. In: Chapman CR, Loeser JD, eds. *Advances in pain research and therapy*. Vol 12. *Issues in pain management*. New York: Raven Press, 1989:391-403.
20 Larue F, Colleau SM, Brasseur L, Cleeland CS. Multicentre study of cancer pain and its treatment in France. *BMJ* 1995;310:1034-7.
21 Serlin RC, Mendoza TR, Nakamura Y, Edwards KR Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. *Pain* 1995;61:277-84.
22 Centers for Disease Control. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR* 1992;41(RR-17):1-19.
23 Cleeland CS, Gonin R, Hatfield AK, Edmonson JH, Blum RH, Steward JA, *et al*. Pain and its treatment in outpatients with metastatic cancer. *N Engl J Med* 1994;330:592-6.
24 Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic research: principles and quantitative methods*. Belmont, California: Lifetime Learning Publications, 1982.
25 Ministry of Social Affairs. *Health care and AIDS. Key figures*. 3rd ed. Paris: Ministère des Affaires Sociales de la Santé et de la Ville, 1994.
26 Vainio A. Treatment of terminal cancer pain in France: a questionnaire study. *Pain* 1995;62:155-62.
27 Larue F, Fontaine A, Brasseur L, Neuwirth L. France: status of cancer pain and palliative care. *J Pain Symptom Manage* 1996;12:106-8.

(Accepted 16 October 1996)

WHEN I USE A WORD . . .

Got an ology?

There are about 400 different -ologies listed in *Chambers Dictionary*, from acar- to zym-, from o- to psychoneuroimmun-. The Greek word logos, from legein meaning to say or read, meant a word, either a word used to express a thought or, by extension, the thought itself. An apologos, literally something that comes from speaking, was a tale, and an apologia was a defence, which is what Plato's *Apology* was: an account of Socrates' defence of himself. "I apologise," my daughter once told me; "I'm not sorry—I apologise."

There are a few English words in which -ology refers to the simple meaning, a word: battology, futile repetition in writing or speech; cacology, bad choice of words or faulty pronunciation; dittology, a double reading or interpretation. In one case it means a syllable: haplology is the loss of a syllable from a word because of the similar sound of a neighbouring syllable; so, mineralology becomes mineralogy, and mammalology becomes mammalogy.

But in about 90% of cases -ology means the study, theory, art, or science of a subject: dactyliology, the study of finger rings; deltiology, picture postcards; vexillology, flags. There is even an -ology for the Turin shroud: sindonology. Unfortunately, there is no ololology (the study of β blockers).

In some cases the -ology has lost its function. Aetiology, for example, should mean the study of the causes of things. However, its original meaning was the assignment of a cause or reason or the reason itself. Only later did it come to mean the science or philosophy of causation. Nowadays it just means "cause," which is the word I prefer to use. Similarly, symptomatology had two roughly contemporaneous meanings, dating from about the turn of the 18th century: the study of symptoms and the symptoms themselves; nowadays it is used to mean just "symptoms." In contrast, methodology originally meant the science of methods, but nowadays its meaning has weakened and it is often used to mean just "methods." I estimate that about 90% of the "methodologies" and "symptomatologies" that you read in scientific papers could be replaced by "methods" and "symptoms." Why don't we use the simpler words?

Jeff Aronson *is a clinical pharmacologist in Oxford*

Randomised comparison of diets for maintaining obese subjects' weight after major weight loss: ad lib, low fat, high carbohydrate diet *v* fixed energy intake

Søren Toubro, Arne Astrup

Abstract

Objectives: To compare importance of rate of initial weight loss for long term outcome in obese patients and to compare efficacy of two different weight maintenance programmes.

Design: Subjects were randomised to either rapid or slow initial weight loss. Completing patients were re-randomised to one year weight maintenance programme of ad lib diet or fixed energy intake diet. Patients were followed up one year later.

Setting: University research department in Copenhagen, Denmark.

Subjects: 43 (41 women) obese adults (body mass index 27-40) who were otherwise healthy living in or around Copenhagen.

Interventions: 8 weeks of low energy diet (2 MJ/day) or 17 weeks of conventional diet (5 MJ/day), both supported by an anorectic compound (ephedrine 20 mg and caffeine 200 mg thrice daily); one year weight maintenance programme of ad lib, low fat, high carbohydrate diet or fixed energy intake diet ($\leq$7.8 MJ/day), both with reinforcement sessions 2-3 times monthly.

Main outcome measures: Mean initial weight loss and proportion of patients maintaining a weight loss of >5 kg at follow up.

Results: Mean initial weight loss was 12.6 kg (95% confidence interval 10.9 to 14.3 kg) in rapid weight loss group and 12.6 (9.9 to 15.3) kg in conventional diet group. Rate of initial weight loss had no effect on weight maintenance after 6 or 12 months of weight maintenance or at follow up. After weight maintenance programme, the ad lib group had maintained 13.2 (8.1 to 18.3) kg of the initial weight loss of 13.5 (11.4 to 15.5) kg, and the fixed energy intake group had maintained 9.7 (6.1 to 13.3) kg of the initial 13.8 (11.8 to 15.7) kg weight loss (group difference 3.5 (-2.4 to 9.3) kg). Regained weight at follow up was greater in fixed energy intake group than in ad lib group (11.3 (7.1 to 15.5) kg *v* 5.4 (2.3 to 8.6) kg, group difference 5.9 (0.7 to 11. 1) kg, $P<0.03$). At follow up, 65% of ad lib group and 40% of fixed energy intake group had maintained a weight loss of >5 kg ($P<0.07$).

Conclusion: Ad lib, low fat, high carbohydrate diet was superior to fixed energy intake for maintaining weight after a major weight loss. The rate of the initial weight loss did not influence long term outcome.

Introduction

Obesity has reached epidemic proportions in the Western world; in Britain the prevalence of obesity doubled between 1980 and 1991, and it is still increasing.[1] Obesity is responsible for a considerable proportion of the mortality attributable to circulatory disease, non-insulin dependent diabetes, and certain cancers, and it is one of the most important avoidable risk factors. In the United States obesity has been estimated to contribute 8% of all health costs.[2] There is strong evidence to suggest that weight loss reverses almost all the health hazards of obesity[3] and normalises mortality.[4] Obese patients inevitably lose weight when they keep strictly to an energy restricted diet,[5] but long term results of energy restriction (the so called calorie counting method) are modest: Typically, after a six month hypocaloric diet (5 MJ/day) has caused 96% of subjects to lose >5 kg weight, only 52% will have maintained a weight loss of ≥ 5 kg at one year's follow up and only 11% will have done so at five years.[6]

Uncertainty over the aetiology of obesity remains a major barrier to developing effective strategies for prevention and treatment. Recent evidence indicates that obesity develops when individuals from the relatively large proportion of the population with a genetic predisposition to obesity are exposed to certain environmental and behavioural conditions, such as an inactive lifestyle and an energy dense diet. A high fat diet seems to promote energy intake by overriding normal signals of satiety,[7 8] and studies from various research disciplines indicate that a high fat diet plays a crucial role in the development and maintenance of obesity.[9] A high fat diet is a risk factor for weight gain among overweight subjects with a family history of obesity,[10] and obese subjects report a habitual diet with a higher fat content than do subjects of normal weight.[11 12]

The apparent role of dietary fat in obesity has lead to the evaluation of ad lib, low fat, high carbohydrate diets for weight loss, but in a short intervention study the weight loss was found to be smaller than from calorie counting.[13] However, as it may take years to develop obesity on a high fat diet, so it may also take years to reverse it by dietary changes. In the present study we therefore decided to induce weight loss by traditional energy restriction and subsequently to compare the long term effects of either an ad lib low fat, high carbohydrate diet or fixed energy intake on weight maintenance. We included two different rates of weight loss in the study in order to examine any effect of this on the long term outcome.

Subjects and methods

Subjects

We consecutively recruited 43 obese adults (2 men, 41 women) who were otherwise healthy from the outpatient waiting list of the Research Department of Human Nutrition, Royal Veterinary and Agricultural University, Copenhagen. All had stable weights, with a body mass index (weight (kg)/(height (m))^2) between 27 and 40. There was one man in each of the two treat-

Research Department of Human Nutrition, Royal Veterinary and Agricultural University, Rolighedsvej 30, 1958 Frederiksberg, Copenhagen, Denmark

Søren Toubro, *associate professor*

Arne Astrup, *head of department*

Correspondence and reprint requests to: Professor Toubro.

BMJ 1997;314:29–34

ment arms, during both weight reduction and weight maintenance. None of the subjects had clinical or biochemical evidence of diabetes or other endocrine disorders or hepatic or renal disease, and none was taking prescribed drugs. They were screened by analysis of blood samples—including haematology (haemoglobin concentration, white cell count, and differential count) and biochemistry (plasma glucose concentration and serum concentrations of sodium, potassium, urate, aspartate aminotransferase, lactate dehydrogenase, alkali phosphatase, and creatinine)—as well as by measurement of blood pressure and electrocardiography. The subjects gave their informed consent to the study according to the declaration of Helsinki II. The protocol was approved by the municipal ethics committee of Copenhagen and Frederiksberg.

Weight reduction programmes

The first phase of the study consisted of weight reduction programmes. All the subjects were randomly assigned to either eight weeks of low energy diet (2 MJ/day) (n = 21) or to 17 weeks of conventional hypocaloric, high protein diet (5 MJ/day) (n = 22). Both diets were supported by an anorectic compound (ephedrine 20 mg and caffeine 200 mg thrice daily,[14] Let-igen, Nycomed DAK A/S, Roskilde, Denmark). The duration of the diets was chosen to achieve similar weight losses in both groups. During this phase, the subjects were split into four groups of 7-13 subjects each (two of low energy diet and two of conventional diet).

The low energy diet consisted of nutrition powder (Bli-Let, Nycomed DAK A/S) dissolved in water and was taken as five daily meals (six for men). This met all recommendations for daily intake of high quality protein (women 60 g, men 72 g), essential amino acids, carbohydrate (30.5 g), vegetable fat (6 g), and fibre (17.5 g). The subjects took a daily supplement of a tablet containing vitamins, minerals, and trace elements and a 1 g fish oil capsule containing at least 350 mg essential omega-3 fatty acids to ensure that their daily intake met recommended amounts.[15]

The conventional diet consisted of ordinary foods and a daily vitamin and mineral tablet (Apovit, Nycomed DAK A/S). The diet plan included recommended amounts of listed food items, together with recipes for low energy marinades and dressings.

The subjects attended the department weekly as outpatients, and the two low energy diet groups were kept apart from the two conventional diet groups. All the groups received nutritional instruction and behaviour therapy[16]: they were instructed by the same staff in dietary guidelines, basic nutritional education, and behaviour therapy sessions of 1-2 hours. Every fortnight the patients were interviewed about adverse effects, and their body weight was measured on a decimal scale (Seca model 707, Copenhagen). Body composition was estimated by the bioimpedance method with an Animeter (HTS-Engineering, Odense, Denmark), and fat free mass and fat mass were calculated with Danish standard equations.[17]

Weight maintenance programmes

After the weight reduction phase the use of low energy diet and anorectic compounds was stopped and subsequently forbidden. The 37 subjects who completed the weight reduction phase were re-randomised to two different one year, non-pharmacological, weight maintenance diet programmes (the low energy diet groups and conventional diet groups being randomised separately). Two groups of patients were assigned to a low fat, high carbohydrate diet consumed ad lib and two groups to a diet of fixed energy intake.

Ad lib, low fat, high carbohydrate diet—The subjects were given a 24 page (A5 format) dietary leaflet specifying details to be included in their daily habits: (a) use a thin layer of butter or margarine on bread or none at all, (b) use the frying pan less often and throw the dripping away, (c) use cooking methods that require less fat, (d) select lean meat and meat products (< 10 g fat/100 g of food item), and (e) eat more carbohydrates, especially complex carbohydrates. Alcoholic beverages were allowed only on special occasions. The aim of the diet was to achieve a macronutrient composition that produced 20-25% of energy intake from fat, at least 55% of energy intake from carbohydrate, and the rest from protein.

Fixed energy intake diet—In order to achieve quantitative as well as qualitative self control, the subjects were introduced to an educational system, which consisted of isoenergetic interchangeable units, represented by 144 counters, each with a small picture of the food it symbolises.[18] The counters were also colour coded, with each colour representing a food group: blue counters for foods rich in protein, green counters for foods rich in fibres and low in energy content, and red counters for foods rich in fat and sugar. The energy content of the food represented by each counter was 260 kJ (62.5 kcal). The subjects were encouraged to restrict the number of red counters they used and to use at least seven blue counters a day. At the start of the programme, the subjects had a daily "ration" of 30 counters (7.8 MJ/day), which could be subsequently adjusted: if a subject's weight increased the daily ration of counters was reduced stepwise by two to a minimum of 20. At a ration of 20 counters, subjects were ordered to complete a seven day record of the weight of their food intake. Once a month, the subjects' body weight, body composition, and any side effects were recorded.

During this phase the subjects were allowed, but not encouraged, to loose weight. They were seen in groups two or three times a month for the first six months, and once a month for the following six months. The sessions consisted of dietary instruction, reinforcement, support, and nutritional education and practical instruction in food preparation in the department kitchen.

One year follow up

At the end of the weight maintenance programme, the 34 subjects who completed it were not contacted again for a year. When they were invited to a follow up visit 28 of these subjects accepted.

Blood analyses

Venous blood samples were drawn from an antecubital vein of each subject after an overnight fast at entry to the study; at the end of the weight reduction phase (after 8 or 17 weeks); after three, six, and 12 months of the weight maintenance phase; and at the one year follow up. Concentrations of plasma glucose, total cholesterol, high density lipoprotein cholesterol, and

Table 1 Anthropometry for all 43 obese subjects entering weight reduction phase of study (values are means (95% confidence interval) unless stated otherwise)

	Weight reduction diet	
	Low energy (8 weeks)	Conventional (17 weeks)
Sex ratio (women:men)	20:1	21:1
Age (years)	43.4 (39.8 to 47.0)	43.8 (39.8 to 47.7)
Height (cm)	166 (163 to 169)	167 (163 to 170)
Weight (kg)	99.5 (93.3 to 105.7)	97.3 (91.9 to 102.7)
Fat free mass (kg)	55.6 (52.4 to 58.8)	54.9 (51.7 to 58.1)
Fat mass (kg)	43.9 (40.1 to 47.6)	42.4 (39.0 to 45.8)
Body mass index (kg/m^2)	36.0 (34.2 to 37.7)	35.0 (33.8 to 36.1)
Weight loss	12.6 (10.9 to 14.3)	12.6 (9.9 to 15.3)

triglycerides were determined by enzymatic methods (Hoffmann la Roche, Basel, Switzerland; Boehringer Mannheim, Mannheim, Germany), while insulin concentrations were determined by radioimmunoassay kits (Novo, Copenhagen, Denmark).

Statistics

Unless stated otherwise, all results are expressed as means (95% confidence interval). Differences between visits and comparisons between different groups were done by either one way or two way analysis of variance (SigmaStat Version 1.02, Jandel Scientific, Germany). All analyses were based on the principle of intention to treat—that is, last body weight was carried forward for the few patients who dropped out. Only the 37 subjects who entered the weight maintenance phase were included in the comparison of the effect of the weight reduction programmes on weight maintenance.

Results

Weight reduction phase

Table 1 shows the subjects' anthropometric data at the start of this phase and the weight loss achieved during this phase. The mean weight loss in the low energy diet groups (12.6 kg (95% confidence interval 10.9 to 14.3 kg)) was similar to that in the conventional diet groups (12.6 (9.9 to 15.3) kg). The rate of weight loss in the low energy diet groups was about twice that in the conventional diet groups (1.6 (1.4 to 1.8) kg/week *v* 0.8 (0.7 to 1.0) kg/week) (see fig 1).

Five subjects withdrew during this phase, two from the low energy diet groups (one refused to drink the nutrition powder and one had an alcohol problem) and three from the conventional diet groups (two due to lack of weight loss and one due to side effects from the anorectic compound). One subject from the low energy diet groups dropped out between the weight reduction and maintenance phases because of psychological problems after involvement in a car accident. The mean weight loss of the 37 subjects who entered the weight maintenance phase was 13.0 (11.4 to 14.6) kg in the low energy diet groups (n = 18) and 14.2 (11.9 to 16.5) kg in the conventional diet groups (n = 19) (group difference 1.2 (– 1.5 to 3.9) kg).

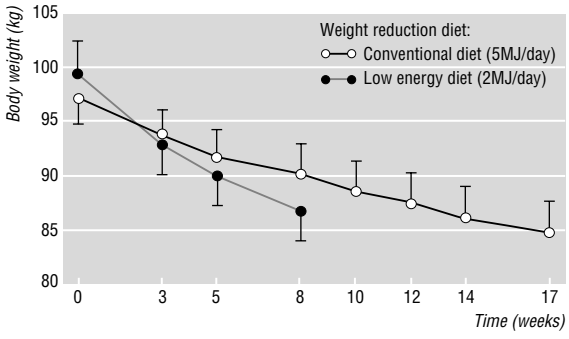

Fig 1 Mean (SE) body weight of subjects during weight reduction phase by type of diet. (Last recorded body weight for dropouts is carried forward)

Weight maintenance phase

The 37 patients were re-randomised to the weight maintenance programmes, producing two similar groups with respect to anthropometry and weight loss. The two groups randomised to the ad lib, low fat diet had lost 13.5 (11.4 to 15.5) kg, and the two groups with fixed energy intake groups had lost 13.8 (11.8 to 15.7) kg. Table 2 shows the subjects' anthropometric data at the start of this phase and the weight regained after two years.

Table 2 Anthropometry, initial weight loss, and weight gain at 1 year follow up for all 37 patients entering weight maintenance phase of study (values are means (95% confidence interval) unless stated otherwise)

	Weight maintenance diet	
	Ad lib, low fat, high carbohydrate	Fixed energy intake
Sex ratio (women:men)	16:1	19:1
Age (years)	43.3 (39.0 to 47.6)	44.6 (40.7 to 48.4)
Height (cm)	167 (164 to 171)	165 (162 to 169)
Weight (kg)	85.9 (78.3 to 93.5)	83.8 (78.6 to 89.0)
Body mass index (kg/m^2)	30.5 (28.6 to 32.3)	30.6 (28.9 to 32.2)
Fat free mass (kg)	52.3 (48.4 to 56.3)	50.7 (47.8 to 53.7)
Fat mass (kg)	33.7 (29.1 to 38.2)	33.2 (29.7 to 36.7)
Initial weight loss (kg)	13.5 (11.4 to 15.5)	13.8 (11.8 to 15.7)
Weight regained at follow up (kg)	5.4 (2.3 to 8.6)	11.3 (7.1 to 15.5)

After six months of the weight maintenance phase the groups with the ad lib diet showed an additional weight loss of 2.3 (0.0 4.6) kg ($P < 0.05$) while the groups with fixed energy intake had gained 0.5 (– 2.6 to 3.6) kg (group difference 2.8 (– 1.0 to 6.6) kg). During the second six months of this phase three patients dropped out—two from the fixed energy intake groups (both due to non-compliance and unavailability to follow up) and one from the ad lib groups (due to development of familial muscle dystrophy). In accordance with analyses being based on the principle of intention to treat, these subjects' last recorded measurements were carried forward. After one year of weight maintenance the ad lib groups showed a non-significant weight gain of 0.3 (– 3.0 to 3.6) kg, while the fixed energy intake groups gained 4.1 (1.2 to 6.9) kg ($P < 0.01$) (group difference 3.8 (– 0.4 to 8.0) kg, $P = 0.08$).

The dietary records carried out three and six months after the start of the weight maintenance programmes did not show any difference between groups in energy intake. The reported percentage of energy intake from dietary fat was lower in the ad lib than in the fixed energy intake groups after three months, but their recorded food intake for seven days covered only about 70% of their estimated energy intake.

Follow up

Thirteen (76%) of the subjects from the ad lib groups and 15 (75%) from the fixed energy intake groups were assessed at follow up one year after the end of the weight maintenance phase. When last recorded body weight was carried forward for all nine subjects who had dropped out, the ad lib groups had regained 5.4 (2.3 to 8.6) kg of the 13.5 kg initially lost while the fixed energy intake groups had regained 11.3 (7.1 to 15.5) kg of the initial 13.8 kg weight loss (group difference 5.9 (0.7 to 11.1) kg, $P<0.03$). The maintained weight loss was also greater in the ad lib groups than in the fixed energy intake groups (8.0 (3.5 to 12.6) kg *v* 2.5 (−1.7 to 6.6) kg, group difference 5.6 (−0.3 to 11.5) kg, $P=0.06$).

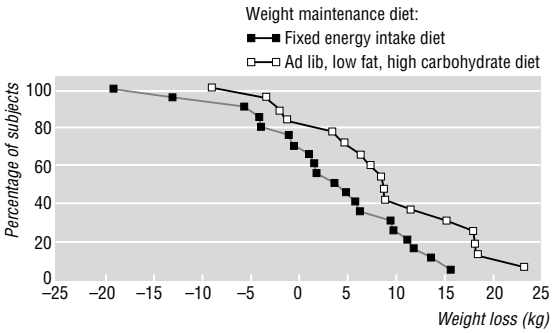

Fig 3 Proportion of subjects who had achieved a given weight loss at follow up by weight maintenance diet

Whole study period

Figure 2 shows the subjects' absolute weight changes. Among the subjects who completed the study, fat mass contributed 67% (3.9 kg) of the total weight regained by those in the ad lib groups compared with 75% (8.5 kg) of the total weight regained by those in the fixed energy intake groups (group difference 3.4 (−1.3 to 8.1) kg). For all subjects who entered the weight maintenance phase, 65% of those in the ad lib groups and 40% of those in the fixed energy intake groups maintained a weight loss >5 kg ($P<0.07$); for the subjects who completed the study, these percentages were 58% and 25% respectively ($P<0.07$).

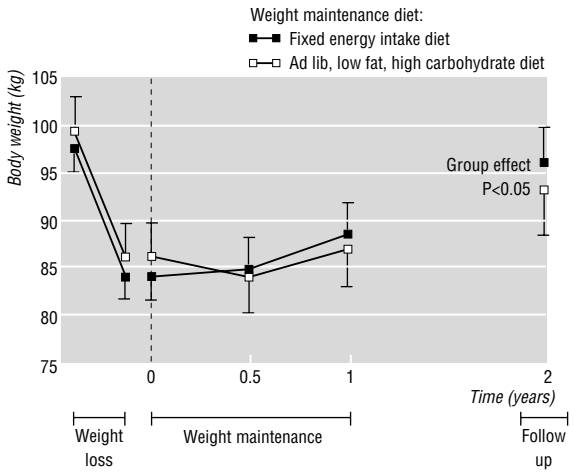

Fig 2 Mean (SE) body weight during weight reduction and weight maintenance phases and at follow up by weight maintenance diet. (Last recorded body weight for drop outs is carried forward)

Figure 3 shows the proportion of subjects for a given weight loss obtained at follow up. The relative risk of subjects in the fixed energy intake groups of maintaining a weight loss <5 kg was 1.7 (0.8 to 3.7) compared with the ad lib groups ($P<0.05$). Analysis of covariance showed that maintenance of weight loss was unaffected by the preceding weight reduction programme (low energy diet or conventional diet) ($P>0.30$); this was true after six months' weight maintenance (group difference 2.4 (−3.2 to 7.9) kg), 12 months' weight maintenance (group difference 3.0 (−2.8 to 8.9) kg), and after one year follow up (group difference 3.0 (−3.0 to 9.1) kg).

We found no group differences in any of the plasma concentrations (glucose, insulin, cholesterol, and triglycerides) measured for blood analyses (see table 3).

Table 3 Plasma concentrations of subjects' blood variables during whole trial period by weight maintenance diet (ad lib, low fat, high carbohydrate diet and fixed energy intake diet)

			Weight maintenance phase			
	At baseline	At end of weight reduction phase	At 3 months	At 6 months	At end (1 year)	At 1 year follow up
Glucose (mmol/l)						
Ad lib, low fat diet	5.0 (4.7 to 5.2)	4.9 (4.6 to 5.1)	4.7 (4.5 to 5.0)	5.1 (4.8 to 5.5)*	4.9 (4.6 to 5.2)	5.0 (4.7 to 5.3)**
Fixed energy intake diet	5.2 (4.9 to 5.5)	5.0 (4.8 to 5.3)	4.9 (4.7 to 5.1)	5.1 (4.8 to 5.4)	4.9 (4.7 to 5.2)	5.2 (4.8 to 5.6)
Insulin (pmol/l)						
Ad lib, low fat diet	98 (69 to 128)	104 (81 to 126)	92 (68 to 117)	118 (89 to 146)**	89 (60 to 119)	75 (43 to 107)**
Fixed energy intake diet	94 (78 to 109)	81 (62 to 99)	89 (71 to 108)	106 (88 to 124)	78 (55 to 101)	68 (44 to 92)
High density lipoprotein cholesterol (mmol/l)						
Ad lib, low fat diet	1.3 (1.1 to 1.4)	1.3 (1.1 to 1.4)	1.3 (1.1 to 1.4)	1.4 (1.2 to 1.5)**	1.4 (1.2 to 1.6)**	1.4 (1.2 to 1.6)**
Fixed energy intake diet	1.2 (1.1 to 1.4)	1.1 (1.0 to 1.3)	1.2 (1.1 to 1.3)	1.3 (1.1 to 1.4)	1.3 (1.2 to 1.5)	1.3 (1.2 to 1.4)
Total cholesterol (mmol/l)						
Ad lib, low fat diet	5.6 (5.0 to 6.2)	5.0 (4.2 to 5.6)	5.3 (4.6 to 6.0)**	5.1 (4.5 to 5.7)*	5.2 (4.6 to 5.8)**	5.7 (5.0 to 6.4)**
Fixed energy intake diet	5.7 (5.1 to 6.2)	4.9 (4.4 to 5.4)	5.3 (4.7 to 5.8)	5.1 (4.6 to 5.7)	5.4 (4.8 to 5.9)	5.7 (5.1 to 6.3)
Ratio of high density lipoprotein cholesterol to total cholesterol						
Ad lib, low fat diet	0.24 (0.20 to 0.28)	0.27 (0.21 to 0.33)	0.25 (0.21 to 0.28)	0.28 (0.23 to 0.32)	0.28 (0.22 to 0.35)	0.25 (0.18 to 0.32)
Fixed energy intake diet	0.22 (0.19 to 0.26)	0.24 (0.31 to 0.27)	0.23 (0.21 to 0.26)	0.26 (0.22 to 0.29)	0.26 (0.22 to 0.30)	0.23 (0.20 to 0.25)
Triglyceride (mmol/l)						
Ad lib, low fat diet	1.6 (1.2 to 2.0)	1.1 (0.9 to 1.3)	1.2 (1.0 to 1.5)	1.2 (1.0 to 1.4)	1.3 (1.0 to 1.6)**	1.5 (1.1 to 1.9)**
Fixed energy intake diet	1.8 (1.3 to 2.3)	1.1 (0.9 to 1.3)	1.2 (1.0 to 1.4)	1.2 (1.0 to 1.5)	1.3 (1.0 to 1.5)	1.5 (1.1 to 1.9)

Student's *t* test for value at end of weight reduction phase versus values at other times: *P<0.05, **P<0.01.

Discussion

Our study shows that a low fat, high carbohydrate diet consumed ad lib was superior to a more traditional fixed energy intake diet in maintaining weight and preventing relapse two years after a major weight loss. The mean weight loss after two years was three times higher in the ad lib group than in the fixed energy intake group (8.0 kg *v* 2.5 kg), and more patients in the former group maintained a substantial proportion of their initial weight loss (maintained weight loss of >5 kg: 65% *v* 40%).

We found that the rate of initial weight loss had no effect on subsequent weight maintenance, which suggests that different procedures to induce weight loss may be equally suitable providing they are followed by an effective, long term dietary programme of weight maintenance. However, with a less intensive weight maintenance programme than the one in this study, we would anticipate the long term outcome after an initial weight loss to be unsatisfactory.

We found no persistent changes in the plasma variables after weight loss, except for plasma insulin concentration, which was significantly reduced in both groups. Hypertriglyceridaemia has been reported in patients with non-insulin dependent diabetes when they were given a high carbohydrate diet[19]; we did not find this in our study, perhaps because of the simultaneous weight loss in our obese subjects who were otherwise healthy.

Role of fat intake in weight maintenance

The mechanism by which a low fat, high carbohydrate diet prevents relapse to a positive energy balance and weight regain may be both by reducing energy intake and by increasing energy expenditure,[12] although the effect on appetite seems to be predominant.[20] In observational studies levels of body fat and obesity are positively associated with dietary fat content,[9] and case reference studies indicate that obese subjects generally consume a diet with a higher fat content than do lean counterparts.[11 21] Although obese subjects tend to underreport their energy intake[22] and possibly also fat intake,[23] when fat oxidation is used as a biological marker of fat intake a larger proportion of dietary energy seems to be derived from fat in obese subjects than in subjects with normal weight.[24]

Meal test studies and short term experimental studies indicate that the satiating effect of fat is weaker than that of carbohydrate[25] and that passive overconsumption of energy can easily be provoked by covertly increasing the fat content of the diet.[7 8] While a high fat diet is not necessarily a prerequisite to the development of obesity, it may promote obesity in subjects with a familial background of obesity[26] or with low levels of physical activity.[1 26] The principle of fixed energy intake is based on voluntary restriction of energy intake, which may tend to keep the dietary composition unchanged—that is, high in fat—whereas a low fat, high carbohydrate diet probably provides greater satiety for less energy intake.

Comparison with other studies

Previous studies have compared the efficacy of ad lib low fat diets with fixed energy intake or calorie counting to induce weight loss, and results for ad lib low fat diets have not been encouraging. Jeffery *et al* found that dietary counselling focusing on fat reduction (20 g/day) and unrestricted carbohydrate intake tended to be more effective than energy restriction to 4.2-5.0 MJ/day[27]: after 12 months' treatment, the low fat group had lost 2.1 kg and the calorie counting group only 0.5 kg. These small weight losses indicate that compliance, at least to the calorie counting, was poor. Though the low fat diet was inefficient in inducing weight loss, the authors found that patients in this group were more compliant with treatment directions, found the diet more palatable, and reported greater reduction in binge eating. In contrast, Schlundt *et al* found that calorie counting produced a larger weight loss in obese subjects than did a low fat ad lib diet during a 16-20 week programme[13]: weight loss in the low fat group was 8.0 kg in men and 3.9 kg in women compared with 11.8 kg in men and 8.2 kg in women in the calorie counting group.

Key messages

- Obesity has reached epidemic proportions in the Western world, but weight loss reverses almost all the health hazards of obesity
- Obese patients lose weight when they keep strictly to an energy restricted diet, but weight losses tend not to be maintained in the long term
- We conducted an intensive one year weight maintenance programme (comparing an ad lib, low fat, high carbohydrate diet with a fixed energy intake diet) after a major weight loss (eight weeks of low energy diet or 17 weeks of conventional diet)
- The rate of initial weight loss did not influence the long term outcome
- The ad lib, low fat diet was superior in maintaining weight loss during weight maintenance programme and at one year follow up

Conclusion

Several points must be considered when comparing the effectiveness of different diets. Firstly, energy restriction as a tool for inducing weight loss is highly effective in obese subjects, and trials have reported mean weight losses of 10-15 kg over six months of treatment.[5 14 16] The relapse rate, however, is remarkably high—50% of all patients regain or exceed their pretreatment weight at 12 months' follow up.[6] Because a high fat content of the diet plays a role for the development of obesity, the low fat principle seems more appropriate as a tool for weight maintenance. We therefore chose to test the two diet principles after an initial weight loss of 12-13 kg.

Secondly, compliance is a crucial factor for weight loss. During their trial of ad lib, low fat diets, Lyon *et al* gave subjects an evening meal to be consumed at home. The meal was enriched with radiolabelled glucose, and the subjects were asked to collect expired air in a test tube after the meal. The recovery of radiolabelled carbon was used as an index of

adherence to the diet, and the authors found a strong correlation between adherence and loss of body fat ($r = 0.74$).[28] Consequently, a small or no weight loss may be attributed to lack of compliance to the diet.

Finally, for nutritional public health policy, our results support the theory that a low fat diet could contribute to preventing obesity. Even a mean weight loss of 2-3 kg may produce a substantial reduction in the prevalence of obesity on a population basis. The recent observation that the prevalence of obesity is increasing concomitantly with a decrease in dietary fat content does not in conflict with this, because the level of daily physical activity is also decreasing.[1] There is an important interaction between exercise level and fat oxidation,[29] indicating that the amount of fat in the diet has to be reduced for a given reduction in total energy expenditure in order to achieve energy balance.

We thank all subjects who participated and the staff of the metabolic unit of the Research Department of Human Nutrition—Lis Kristoffersen, Tina Cuthbertson, Charlotte Kostecki, Inge Timmermann, Bente Knap.

Funding: Drugs and part of the study's funding were provided by Nycomed DAK A/S, Roskilde, Denmark.

Conflicts of interest: None.

1 Prentice AM. Obesity in Britain: gluttony or sloth. *BMJ* 1995;311:437-9.
2 Colditz GA. Economic costs of obesity. *Am J Clin Nutr* 1992;55:503-7S.
3 Garrow J. Importance of obesity. *BMJ* 1991;303:704-6.
4 Williamson DF, Pamuk E, Thun M, Flanders D, Byers T, Heath C. Prospective study of intentional weight loss and mortality in never-smoking overweight US white women aged 40-64 years. *Am J Epidemiol* 1995;141:1128-41.
5 Garrow JS, Webster JD. Effects on weight and metabolic rate of obese woman of a 3.4 MJ (800 kcal) diet. *Lancet* 1989;i:1429-31.
6 Wadden TA, Sternberg JA, Letizia KA, Stunkard AJ, Foster GA. Treatment of obesity by very low calorie diet, behaviour therapy, and their combination: a five-year perspective. *Int J Obes* 1989;13(suppl 2):39-46.
7 Stubbs JR, Ritz P, Coward WA, Prentice AM. Covert manipulation of the ratio of the dietary fat to carbohydrate and energy density: effect on food intake and energy balance in free-living men eating ad libitum. *Am J Clin Nutr* 1995;62:330-7.
8 Stubbs JR, Harbron CG, Murgatroyd PR, Prentice AM. Covert manipulation of dietary fat and energy density: effect on substrate flux and food intake in men eating ad libitum. *Am J Clin Nutr* 1995;62:316-29.
9 Lissner L, Heitmann BL. Dietary fat and obesity: evidence from epidemiology. *Eur J Clin Nutr* 1995;49:79-90.
10 Heitmann BL, Lissner L, Sørensen TIA, Bengtsson C. Dietary fat intake and weight gain in women genetically predisposed to obesity. *Am J Clin Nutr* 1995;61:1213-7.
11 Dreon DM, Frey-Hewitt B, Ellsworth N, William PT, Terry RB, Wood PD. Dietary fat carbohydrate ratio and obesity in middle-aged men. *Am J Clin Nutr* 1988;47:995-1000.
12 Astrup A, Raben A. Carbohydrate and obesity. *Int J Obes* 1995; 19(suppl 5):S27-37.
13 Schlundt DG, Hill JO, Pope-Cordle J, Arnold D, Virts KL, Katahn M. Randomised evaluation of low fat ad libitum carbohydrate diet for weight reduction. *Int J Obes* 1993;17:623-9.
14 Astrup A, Breum L, Toubro S, Hein P, Quaade F. The effect and safety of an ephedrine/caffeine compound to ephedrine, caffeine and placebo in obese subjects on a energy restricted diet. A double blind trial. *Int J Obes* 1992;16:269-77.
15 Hoie LH, Brusgaard D, Thom E. Reduction of body mass and change in body composition on a very low calorie diet. *Int J Obes* 1993;17:17-20.
16 Stunkard AJ, Wadden TA. *Obesity, theory and therapy.* 2nd ed. New York: Raven Press, 1993: 197-219.
17 Heitmann BL. Prediction of body water and fat in adult Danes from measurement of electrical impedance. A validation study. *Int J Obes* 1990;14:789-802.
18 Hey H, Pedersen HD, Andersen T, Quaade F. Formula diet plus free additional food choice up to 1000 Kcal(4.2 MJ) compared with an isoenergetic conventional diet in the treatment of obesity. A randomised clinical trial. *Clin Nutr* 1987;6:195-9.
19 Chen IY-D, Coulston AM, Zhou MY, Hollenbeck C, Reaven GM. Why do low-fat high carbohydrate diets accentuate postprandial lipemia in patients with NIDDM? *Diabetes Care* 1995;18:10-6.
20 Hill JO, Prentice AM. Sugar and body weight regulation. *Am J Clin Nutr* 1995;62(suppl 1):S264-74.
21 Tucker LA, Kano MJ. Dietary fat and body fat: a multivariate study of 205 adult females. *Am J Clin Nutr* 1992;56:616-22.
22 Prentice AM, Black AK, Coward WA, Davis HL, Goldberg GR, Murgatroyd PR, *et al.* High levels of energy expenditure in obese women. *BMJ* 1986;292:983-7.
23 Heitmann BL, Lissner L. Dietary under reporting by obese individuals—is it specific or nonspecific? *BMJ* 1995;311:986-9.
24 Astrup A, Buemann B, Western P, Toubro S, Raben A, Christensen NJ. Obesity as an adaptation to a high-fat diet: evidence from a cross-sectional study. *Am J Clin Nutr* 1994;59:350-5.
25 Blundell JE, Burley VJ, Cotton JR, Lawton CL. Dietary fat and the control of energy intake: evaluating the effects of fat on meal size and postmeal satiety. *Am J Clin Nutr* 1993;57(suppl):S772-8.
26 Trembly A. Difference in fat balance underlying obesity. *Int J Obes* 1995;19(suppl 7):S10-4.
27 Jeffry RW, Hellerstedt WL, French SA, Baxter JE. A randomized trial of counselling for fat restriction versus calorie restriction in the treatment of obesity. *Int J Obes* 1995;19:132-7.
28 Lyon X-H, Di Vetta V, Milon H, Jequier E, Schutz Y. Compliance to dietary advice directed towards increasing the carbohydrate to fat ratio of the every day diet. *Int J Obes* 1995;19:260-9.
29 Tremblay A, Almeras N, Boer J, Kranenberg EK, Despres J-P. Diet composition and postexercise energy balance. *Am J Clin Nutr* 1994;59:975-9.

(Accepted 16 October)

Longitudinal study of the effect of apolipoprotein e4 allele on the association between education and cognitive decline in elderly men

Sandra Kalmijn, Edith J M Feskens, Lenore J Launer, Daan Kromhout

Department of Chronic Diseases and Environmental Epidemiology, National Institute of Public Health and the Environment, PO Box 1, 3720 BA Bilthoven, Netherlands

BMJ 1997;314:34–5

continued over

Less educated people have an increased risk of cognitive decline,[1] and several possible explanations for this have been suggested. The activation of nerve cells in higher educated subjects might protect these cells against degeneration,[2] thereby delaying the pathological process leading to cognitive decline. Higher education might also lead to an increased brain reserve capacity, so that other neurons can take over the tasks of dead neurons.[3] This would retard the onset of cognitive decline. However, the association might be due to confounding by factors related to lifestyle.

The apolipoprotein e4 allele on chromosome 19 is an important risk factor for cognitive decline.[4] We examined the association between education and cognitive decline in elderly men with and without the e4 allele to see whether this genetic risk factor modified the association.

Subjects, methods, and results

The Zutphen elderly study is a longitudinal study on risk factors for chronic diseases in men living in Zutphen, the Netherlands.[4] In 1990, 560 (78%) of the 718 surviving men were examined, and the examinations were repeated in 1993 on 390 (71%) of the 553 survivors. Complete information was available for 356

men participating both in 1990 and 1993. Global cognitive function was tested with the Dutch version of the 30 point mini-mental state examination. Cognitive decline was defined as a drop of more than two points (>1 standard deviation) during 1990-3, which corresponded to the 14th centile of change. Years of formal education were obtained in 1990 and categorised as ≤6 years or >6 years. Serum samples were obtained in 1990 and frozen at −20°C until 1993, when the apolipoprotein E phenotype was determined by isoelectric focusing of delipidated plasma samples followed by immunoblotting. The association between education and cognitive decline was examined by logistic regression in the total group and the carriers and non-carriers of e4, adjusting for age and baseline score for cognitive function.

The subjects' mean age in 1990 was 74.6 (SD 4.2) years. The median score for cognitive function was 27 (10th centile 23, 90th centile 29). Table 1 shows the association between education and cognitive decline. This was strong in non-carriers of e4, but in carriers of e4 it was absent (test for interaction not significant, P = 0.11). Additional adjustment for history of cardiovascular diseases did not change these results. For carriers and non-carriers of e4, the subjects who did not participate in 1993 showed a similar association between education and cognitive function in 1990 as did those who participated in 1993 (results not shown).

Table 1 Adjusted odds ratios (95% confidence interval) for cognitive decline* in 356 elderly men according to level of education and presence of apolipoprotein e4 allele

Education	Total population (n=356)	Carrying e4 allele	
		No (n=272)	Yes (n=84)
>6 Years (n=313)†	1	1	1
≤6 Years (n=43):			
Adjusted for age	1.8 (0.8 to 4.0)	2.6 (1.0 to 7.1)	0.7 (0.2 to 3.1)
Adjusted for age and baseline cognitive function	2.1 (0.9 to 4.9)	3.1 (1.1 to 8.8)	0.9 (0.2 to 3.8)

*Drop of >2 points on mini-mental state examination between 1990 and 1993.
†Reference value.

Comment

We observed a significantly increased risk of cognitive decline associated with a lower level of education in subjects without an apolipoprotein e4 allele. In contrast, there was no association between education and cognitive decline in carriers of the e4 allele.

Possible explanations for this lack of association in the subjects with an e4 allele include selective survival and non-response. However, our data do not suggest that less educated men with e4 who did not participate in 1993 had a different risk of cognitive impairment than those who participated. We do not know, though, whether the risk of cognitive decline differed between the groups. The e4 allele may be such a strong risk factor for cognitive decline that cognitive performance in carriers of an e4 allele deteriorates regardless of educational level. Alternatively, apolipoprotein e4 may play a role in inhibiting neuronal growth[5] and may thus block the putative stimulating effect of education on neuronal growth. Finally, we cannot exclude the possibility that our results are a chance finding.

In conclusion our results suggest that the apolipoprotein e4 allele may modify the association between education and cognitive decline. However, our findings need to be confirmed by larger studies.

Sandra Kalmijn, *research associate*
Edith J M Feskens, *senior researcher*
Lenore J Launer, *assistant professor*
Daan Kromhout, *professor of public health research*

Correspondence and reprint requests to: Dr Kalmijn (email Kalmijn@epib.fgg.eur.nl).

We thank the fieldwork team in Zutphen, especially Drs E B Bosschieter and B P M Bloemberg; C de Lezenne Coulander for data management; and Dr L M Havekes for determination of the apolipoprotein E polymorphism.

Funding: This study was supported by grants from the Praeventie Fonds, the Hague, Netherlands, and the National Institute on Aging, Bethesda MD, USA.

Conflict of interest: None.

1 White L, Katzman R, Losonczy K, Salive M, Wallace R, Berkman L, *et al.* Association of education with incidence of cognitive impairment in three established populations for epidemiologic studies of the elderly. *J Clin Epidemiol* 1994;47:363-74.
2 Swaab DF. Brain aging and Alzheimer's disease, "wear and tear" versus "use it or lose it." *Neurobiol Aging* 1991;12:317-24.
3 Katzman R. Education and the prevalence of dementia and Alzheimer's disease. *Neurology* 1993;43:13-20.
4 Feskens EJM, Havekes LM, Kalmijn S, de Knijff P, Launer LJ, Kromhout D. Apolipoprotein e4 allele and cognitive decline in elderly men. *BMJ* 1994;309:1202-6.
5 Mahley RW, Nathan BP, Bellosta S, Pitas RE. Apolipoprotein E: impact of cytoskeletal stability in neurons and the relationship to Alzheimer's disease. *Curr Opin Lipidology* 1995;6:86-91.

(Accepted 23 July 1996)

Coma induced by abuse of γ-hydroxybutyrate (GBH or liquid ecstasy): a case report

Gareth Thomas, Stephen Bonner, Alastair Gascoigne

We present a case of coma and respiratory depression secondary to γ-hydroxybutyrate abuse, necessitating mechanical ventilation and intensive therapy. This drug was available as a health food in the United States until it became a drug of abuse. It is now being taken in the United Kingdom.

Case report

A 32 year old man presented from a night club. His Glasgow coma score was 5, with pinpoint pupils and absent laryngeal reflexes. There was no evidence of head injury or intravenous drug abuse. His pulse rate was 30-50 beats/min, blood pressure 100/70 mm Hg, and respiratory rate 4 breaths per minute. His respiratory efforts were assisted with 100% oxygen using a bag and mask. There was no response to intravenous naloxone or flumazenil. His trachea was intubated without the need for muscle relaxants and his lungs ventilated.

A cranial computed tomography scan was normal and the patient was transferred to the intensive care

Intensive Care Unit, Royal Victoria Infirmary, Newcastle upon Tyne NE1 4LP

BMJ 1997;314:35–6

continued over

Gareth Thomas, *senior house officer*
Stephen Bonner, *senior registrar in anaesthesia and intensive therapy*
Alastair Gascoigne, *consultant in intensive therapy and respiratory medicine*
Correspondence to: Dr Bonner.

unit. No paracetamol or salicylates were found in his serum, and toxicological analysis of his urine was negative for amphetamines, barbiturates, cocaine metabolites, methadone, opiates, and alcohol but positive for benzodiazepines. The diagnosis was uncertain but drug ingestion was suspected.

Twelve hours after admission his conscious level rose; he was extubated and made an uneventful recovery, the only adverse effect being myalgia. He admitted having taken a tablet of temazepam and "half of a bottle" of GBH (γ-hydroxybutyrate).

Comment

γ-Hydroxybutyrate is a catabolite of γ-aminobutyric acid; is found in the central nervous system, kidney, heart, and muscle; is freely permeable across the blood brain barrier; and may play a role as a central neurotransmitter. Synthesised in the 1960s, it was used as a sedative premedicant and an intravenous anaesthetic induction agent, of particular value in children,[1] but it fell into disfavour owing to its lack of analgesic properties and epileptogenic side effects.

It is available as a colourless, odourless liquid, powder, or capsules. It may be injected but is usually taken orally. Low doses produce euphoria, but higher doses produce sedation. An oral dose of 10 mg/kg produces amnesia, 30 mg/kg promotes sleep, and doses above 50 mg/kg produce general anaesthesia.[2] Above this dose cardiac output falls, and respiratory depression and seizure-like activity may occur.[2] Since most γ-hydroxybutyrate is illegally manufactured, 40 ml (3-9 doses) may contain a dose as small as 3 g or one as potentially toxic as 20 g.[3]

Clinical effects are potentiated by concurrent sedative drugs, such as ethanol, opiates, benzodiazepines, and neuroleptics.[2] Minor side effects include bradycardia, hypotension, nausea and vomiting, diarrhoea, urinary incontinence, increased libido, dizziness, tremor, headache, and euphoria. More severe effects include myoclonus, ataxia, tunnel vision, confusion, agitation, hallucinations, tonic-clonic seizures, decreased conscious level, respiratory arrest, and coma.[2 4]

γ-Hydroxybutyrate poisoning has been widely reported in the United States but not so far in the United Kingdom. The drug had been marketed in the US by health food stores to promote body building, dieting, and "natural sleep,"[5] but subsequently it became a drug of abuse for its euphoric effects. In 1990 the US Food and Drug Administration reported adverse effects due to the illicit use of γ-hydroxybutyrate (commonly called GBH or liquid ecstasy), and was banned in some states.[5] In three months 57 cases of γ-hydroxybutyrate poisoning were reported in California, nine requiring ventilation or intensive care.[4] In the United Kingdom the drug is not controlled under the Misuse of Drugs Act, so possession is not an offence.[3] It is usually found in the club scene,[3] and anecdotal evidence from the north east of England suggests that the use of the drug is on the increase.

γ-Hydroxybutyrate poisoning should be considered in the differential diagnosis of patients presenting in coma. Diagnostic tests are not widely available; recovery is spontaneous with supportive treatment, which may require mechanical ventilation and intensive care.

Funding: None.
Conflict of interest: None.

1 Hunter AS, Long WJ, Ryrie CG. An evaluation of gamma-hydroxybutyric acid in paediatric practice. *Br J Anaesth* 1971;43:620-7.
2 Multistate outbreak of poisonings associated with illicit use of gamma hydroxy butyrate. *JAMA* 1991;265:447-8.
3 Shapiro H. *Information for drug workers: factsheet No 8 GHB.* London: Institute for the Study of Drug Dependence, 1994.
4 Warning about GHB from the Food and Drug Administration. *JAMA* 1991;265:1802.
5 Chin M, Kreutzer R, Dyer J. Acute poisoning from gamma hydroxybutyrate in California. *West J Med* 1992;156:380-4.

(Accepted 18 July 1996)

Why don't cancer patients get entered into clinical trials? Experience of the Sheffield Lymphoma Group's collaboration in British National Lymphoma Investigation studies

B W Hancock, M Aitken, C Radstone, G Vaughan Hudson

Yorkshire Cancer Research Campaign Department of Clinical Oncology, Weston Park Hospital, Sheffield S10 2SJ

BMJ 1997;314:36–7

continued over

By standardising treatment, often in collaboration with specialised centres, multicentre controlled clinical trials offer higher survival rates, particularly for less common cancers.[1] Patients are generally willing to take part in such studies,[2] but in oncology generally few patients are entered into trials. Since 1981 the Sheffield Lymphoma Group, which sees patients by tertiary referral from district general hospitals in north Trent, has committed itself to entering patients into current British National Lymphoma Investigation studies. We looked at our record in doing so over 12 years.

Methods and results

We looked in detail at all patients referred to us in 1981-92, particularly to determine why patients were not entered into studies.

Of 1927 patients referred, 1813 were accepted by the group as having histologically proved lymphoma, and 822 (45%) of these patients were entered into studies. The annual recruitment rate varied from 25% in 1992 to 61% in 1984 (table 1). Reasons for non-entry were as follows.

Table 1 Numbers of patients referred to and accepted by the Sheffield Lymphoma Group and entered into British National Lymphoma Investigation studies in 1981-92

	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	Total
Referred to Sheffield Lymphoma Group	135	142	144	140	141	164	148	187	191	186	168	181	1927
Accepted by Sheffield Lymphoma Group	128	133	135	137	135	154	148	177	176	170	156	164	1813
Recruited into British National Lymphoma Investigation studies	60	70	73	84	76	78	71	91	65	68	45	41	822

- Forty four per cent (804) were not eligible for current studies because they were too old (274), they were medically ineligible (365), no current study was available (55), they proved not to have lymphoma (73), or there was histological disagreement on their lymphoma type (37).
- Ten per cent (187) were eligible but not entered into studies: 128 were considered but not entered by the clinician, 55 were missed by the clinician, and 4 refused.

Comment

The status of clinical trials has been under the spotlight recently.[3 4] Stiller reviewed published data on survival rates for cancer in relation to patterns of organisation of medical care—specifically treatment by protocol, usually within the context of a clinical trial, and also treatment at specialist centres.[1] Entry into trials (and centralised referral) was often associated with higher survival rates, particularly with less common cancers, and was never found to be associated with a lower survival rate. Patients themselves are usually willing to participate, attracted by being treated by a doctor with a specialist interest and encouraged by the promise of close monitoring of their progress.[2] In our series only four refused to enter when asked.

Many patients are considered for studies but not entered on the clinician's decision. Even with our own group's positive attitude to trials and policy of putting as many patients as possible into studies, this explanation has consistently accounted for around 7% of patients. Almost half of our patients, however, were not considered for trials for genuine reasons: the eligibility criteria of particular trials excluded them on age or medical grounds or there was no appropriate current trial (old trials finished, new trials not yet started—as was the case in the last two years of our study period, when recruitment was temporarily lower). Such selection may be an important factor in determining the different outcomes between patients who take part in trials and those who do not. In particular, older and iller patients tend not to be entered into studies, and survival in this group will naturally be poor.

An increasing number of patients with cancer are eligible for relatively straightforward relevant clinical studies with less strict eligibility criteria, and these patients have the right to be fully informed about,[4] and to have the opportunity to participate in, such trials. Major deterrents to involvement in trials by doctors are lack of time (to gain informed consent, to obtain ethics committee approval, to collect data) and lack of staff (many clinicians, particularly NHS staff in non-teaching hospitals, do not have access to dedicated research staff).[5] Nevertheless, it should still be possible for motivated centres, actively encouraged and kept up to date by enthusiastic study teams, to recruit sizeable numbers of patients into large randomised clinical trials, hopefully to the benefit of the individual and to the disease group as a whole.

We thank the staff at the British National Lymphoma Investigation secretariat, the Trent Cancer Registry for providing details of tumour registrations, and all the clinicians who have cared for patients featured in this study.

Funding: Yorkshire Cancer Research Campaign and Weston Park Hospital Cancer Research Trust.

Conflict of interest: None.

1 Stiller CA. Centralised treatment, entry to trials and survival. *Br J Cancer* 1994;70:352-62.
2 Slevin M, Mossman J, Bowling A, Leonard R, Steward W, Harper P, *et al.* Volunteers or victims: patients' views of randomised cancer clinical trials. *Br J Cancer* 1995;71:1270-4.
3 Tannock IF. The recruitment of patients into clinical trials. *Br J Cancer* 1995;71:1134-5.
4 Smyth JF, Mossman J, Hall R, Hepburn S, Pinkerton R, Richards M, *et al* on behalf of the United Kingdom Co-ordinating Committee on Cancer Research. Conducting clinical research in the new NHS: the model of cancer. *BMJ* 1994;309:457-61.
5 Smith R, Goodare H. The rights of patients in research. *BMJ* 1995; 310:1277-8.
(Accepted 19 September 1996)

B W Hancock, *professor*
M Aitken, *research assistant*
C Radstone, *clinical trials officer*
British National Lymphoma Investigation, Middlesex Hospital, London W1N BAA
G Vaughan Hudson, *director*
Correspondence to: Professor Hancock.

A PATIENT WHO CHANGED MY PRACTICE

The innocuous statement

Following the retirement of my senior partner, who had served the village community for 35 years, patients were now obliged to consult me, a relative newcomer of two and a half years. One such patient was an elderly woman with a moderately severe deformity of the hands due to osteoarthritis. As she was able to walk to the surgery she had been a regular attender. She had been prescribed a full range of the drugs which were available at that time. Although I cannot remember precisely what I said, as I handed her the repeat prescription I told her that her previous doctor had given her the best medication then available and I could not prescribe anything better.

In those days we did not have appointments and ancillary staff—patients came to the open surgery at their convenience. As a result, this patient went out of my mind and I did not see her again for some months when she requested a home visit. To my dismay I found that she was now quite disabled, with most of her joints affected by arthritis. I said that I wished that she had sent for me earlier and her reply was, "When I last saw you you told me that you could not give me anything better, so I saw no point in coming when the tablets were not curing me."

An essential feeling to be given by the doctor to every patient is "hope," and since that day I have tried to choose my words with care.

Eric Hainsworth *is a retired general practitioner living in Cornwall*