Open Access: free full text available online
For the full versions of these articles see bmj.com

RESEARCH

Secular trends in self reported sexual activity and satisfaction in Swedish 70 year olds: cross sectional survey of four populations, 1971-2001

Nils Beckman, Margda Waern, Deborah Gustafson, Ingmar Skoog

EDITORIAL by Kleinplatz

Neuropsychiatric Epidemiology Unit, Institute of Neuroscience and Physiology, Sahlgrenska Academy at Gothenburg University, Gothenburg, Sweden

Correspondence to: N Beckman, Neuropsychiatric Epidemiology Unit, Neuropsykiatri SU/Mölndal, Wallinsgatan 6, SE 431 41 Mölndal
nils.beckman@neuro.gu.se

Cite this as: *BMJ* 2008;337:a279
doi:10.1136/bmj.a279

ABSTRACT

Objective To study secular trends in self reported sexual behaviour among 70 year olds.

Design Cross sectional survey.

Settings Four samples representative of the general population in Gothenburg, Sweden.

Participants 1506 adults (946 women, 560 men) examined in 1971-2, 1976-7, 1992-3, and 2000-1.

Main outcome measures Sexual intercourse, attitudes to sexuality in later life, sexual dysfunctions, and marital satisfaction.

Results From 1971 to 2000 the proportion of 70 year olds reporting sexual intercourse increased among all groups: married men from 52% to 68% (P=0.002), married women from 38% to 56% (P=0.001), unmarried men from 30% to 54% (P=0.016), and unmarried women from 0.8% to 12% (P<0.001). Men and women from later birth cohorts reported higher satisfaction with sexuality, fewer sexual dysfunctions, and more positive attitudes to sexuality in later life than those from earlier birth cohorts. A larger proportion of men (57% *v* 40%, P<0.001) and women (52% *v* 35%, P<0.001) reported very happy relationships in 2000-1 compared with those in 1971-2. Sexual debut before age 20 increased in both sexes: in men from 52% to 77% (P<0.001) and in women from 19% to 64% (P<0.001).

Conclusion Self reported quantity and quality of sexual experiences among Swedish 70 year olds has improved over a 30 year period.

INTRODUCTION

Most elderly participants in surveys on sexual behaviour developed their views during the early part of the 20th century. We examined secular trends in sexual behaviour and attitudes to sexuality in later life in 70 year olds from Gothenburg, Sweden.

METHODS

Sample populations

Four representative samples of 70 year olds from Gothenburg were surveyed in 1976-7, 1992-3, and 2000-1, to study secular trends in health. The samples were obtained from the Swedish population register. No significant differences existed between responders and non-responders (see bmj.com).[1 2]

Sample 1—participants born between 1 July 1901 and 30 June 1902 on dates ending with 2, 5, or 8 were invited to a health examination in 1971-2[3]; 392 of 460 (85.2%) participated in a psychiatric examination.[1]

Sample 2—participants born between 1 July 1906 and 30 June 1907 on dates ending with 2, 5, or 8 were invited to a health examination in 1976-7; 404 of 513 (78.8%) participated in a psychiatric examination.[2]

Sample 3—70 year old women born in 1922 on days 6, 12, 18, 24, or 30 were invited to a health examination in 1992-3; 249 of 381 (65.4%) participated in a psychiatric examination.

Sample 4—70 year olds born in 1930 on days 3, 6, 12, 18, 21, 24, or 30 were invited to a health examination in 2000-1; 500 of 767 (65.2%) participated in a psychiatric examination.

The psychiatric examination included questions on sexual behaviour: attitudes to sexuality in later life, frequency of intercourse during the past year, and age of sexual debut and its timing in relation to marriage. Sexual activity was defined as having had intercourse during the past year. Intercourse was defined as sexual contact, most often with penetration. Questions asked in the examinations of all but the first sample were about whether sexuality was a positive or negative factor in life, satisfaction with intercourse, sexual dysfunction, and reason for cessation of intercourse.

One of the researchers (IS), a psychiatrist, was trained by those who did the examinations in the 1970s, and trained those who did the examinations in 1992 and 2000. Inter-rater agreement on frequency of intercourse and attitudes to sexuality at age 70 were high.

Statistical analysis

We categorised marital status as married or cohabiting compared with unmarried. Educational level was dichotomised as compulsory or more than compulsory.

Differences in proportions were tested using Fisher's exact test. The Cochran-Armitage χ^2 test was used to test for trends. We used an asymptotic permutation test of trend for differences in the median age of sexual debut. Data were analysed by strata of sex and marital status. For regression analyses we also pooled data from all the samples. We used binary logistic regression

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a279

models to estimate the odds of reporting intercourse (yes or no within the past year) by sample (1971-2 plus 1976-7 *v* 1992-3 plus 2000-1), marital status, male gender, sexual debut before age 20, a positive attitude towards sexuality in later life, diagnosis of depression, educational level, and three year mortality. We present the associations as odds ratios and 95% confidence intervals. In all analyses we used two tailed tests. We considered results significant at $P<0.05$.

RESULTS

Among both sexes the proportion of participants who were divorced, cohabiting, or in a relationship but living apart increased over the 30 years of sampling (see bmj.com). Among those who had a partner, the proportion reporting a happy relationship increased in both sexes. Compared with men, women in all samples were less often married or cohabiting, more often widowed, and more often had an older partner.

The proportion of 70 year olds reporting that they were sexually active, that sexuality had been a positive factor in their life, and that had a positive attitude to sexuality in later life increased during the study period, both among married and cohabiting participants and among unmarried participants (table). Fewer people in later cohorts reported never having had intercourse. Among those reporting intercourse, the proportion that had intercourse at least once a week increased over the 30 year period. Concurrently the reported median age of sexual debut decreased in both sexes and the proportion reporting premarital intercourse increased in women. Reported intercourse was more common among men than among women in all four samples, and men reported an earlier age of sexual debut than women although the differences between the sexes for this variable diminished among those from later born samples.

In a logistic regression analysis including the entire sample, being in a later born cohort increased the odds of having intercourse (odds ratio 1.48, 95% confidence interval 1.10 to 2.00), independent of marital status, sex, sexual debut before age 20, a positive attitude to sexuality in later life, depression, educational level, and three year mortality.

The proportion of women reporting high or very high sexual satisfaction increased and reports of no sexual satisfaction decreased from the second to the last samples (see bmj.com).

Among those who had a partner, both sexes reported that in most cases cessation of intercourse was due to male related factors (see bmj.com). This pattern did not change over the 30 year period.

DISCUSSION

Self reported sexual activity among 70 year olds in Gothenburg, Sweden increased from 1971 to 2001. At the same time among elderly people attitudes to sexuality became more positive, and the proportion reporting a very happy relationship increased. Furthermore, the proportion reporting high satisfaction with sexual activity and that sexuality was an important factor in life increased. Consistent with population studies of younger samples of later born cohorts[4-6] the median age of sexual debut decreased and the proportion that had their sexual debut before age 20 increased. The one year prevalence of intercourse in the two earliest birth cohorts was similar to that among septuagenarians reported from studies in the 1950s and

Self reported sexual behaviour and attitudes in four samples of 70 year olds from Gothenburg, Sweden, examined in 1971-2, 1976-7, 1992-3, and 2000-1. Values are number who answered question of total number examined (percentage) unless stated otherwise

	Men				Women				
Variable	1971-2 (n=161)	1976-7 (n=174)	2000 (n=225)	Sample trend P value†	1971-2 (n=221)	1976-7 (n=222)	1992 (n=241)	2000 (n=262)	Sample trend P value†
Positive attitude towards sexuality in old age	121/148 (82)	139/173 (80)	200/207 (97)	0	135/208 (65)	135/215 (63)	212/238 (89)	219/232 (94)	0
Married or cohabiting	97/117 (83)	106/133 (80)	162/168 (96)	0	62/89 (70)*	55/92 (60)**	119/135 (88)	109/117 (93)	0
Not married	24/31 (77)	33/40 (83)	38/39 (97)	0.010	73/119 (61)	80/123 (65)*	93/104 (89)	110/115 (96)	0
Sexuality a positive factor in life	—	44/174 (26)	196/206 (95)	0‡	—	10/214 (5)***	121/225 (54)	181/231 (78)***	0
Sexual intercourse during past year	72/152 (47)	83/173 (48)	133/203 (66)		35/209 (12)	39/213 (18)	81/232 (35)	77/225 (34)	
Married or cohabiting	62/119 (52)*	71/133 (53)*	113/166 (68)	0.002	34/89 (38)***	34/92 (37)***	70/130 (54)***	63/112 (56)***	0
Not married	10/33 (30)	12/40 (30)	20/37 (54)	0.016	1/120 (1)***	5/121 (4)***	11/102 (11)	14/113 (12)***	0
Sexual intercourse once weekly or more among sexually active	7/72 (10)	22/83 (27)	41/133 (31)	0.006	3/35 (9)	7/39 (18)	16/81 (20)	20/77 (26)	0.047
Sexual debut before age 20 (median age at sexual debut)	77/148 (52) (19.3)	94/167 (56) (18.7)	159/207 (77) (17.7)	0§	39/203 (19) (22.7)	57/206 (28) (22.0)	112/230 (49) (19.6)	147/229 (64) (18.6)	0§
Sexual intercourse before marriage	123/149 (83)	147/171 (86)	183/207 (88)	0.151	86/180 (48)***	143/194 (74)**	172/229 (75)	198/226 (88)	0
Sexually inexperienced	1/156 (1)	0/173 (0)	0/207 (0)	0.308	23/210 (11)***	15/212 (7)***	1/231 (0.4)	1/229 (0.4)	0

Number of participants varies within cohorts as some declined to answer some questions.
*P<0.05; **P<0.01; ***P<0.001 (Fisher's exact test) for difference between sexes in birth sample, or difference between participants who were married or cohabiting and not married.
†Cochran-Armitage χ^2 test for sample trend.
‡Fisher's exact test.
§Test for column trend with asymptotic permutation test of trends.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Secular trends in elderly people's sexual behaviour is unclear

WHAT THIS STUDY ADDS

The quantity and quality of sexual experiences among 70 year olds in Sweden improved over a 30 year period

Attitudes to sexuality have become more positive in this age group

1980s.[7 8] The prevalence in the two younger birth cohorts is similar to a European study in 2001-2[9] and a US study in 2005-6.[10]

Whether elderly couples continue to be sexually active seems to a large extent to be determined by men.[11] This pattern, which did not change over time, was also reported in studies in the 1950s[12] and in 2005-6.[10]

In agreement with previous reports self reported sexual activity was more common in men, regardless of marital status.[8-10 12-16] Differences between the sexes in self reported sexual behaviour, however, decreased from 1971 to 2001 among the 70 year olds in our study. Overall, men reported an earlier age of sexual debut and a higher proportion of premarital sex than women in the 1970s, but this sex difference diminished among those in later born samples. Finally, whereas 70 year old men in the 1970s more often reported positive attitudes to sexuality than women, there were no sex differences in attitudes in 2000-1. Thus attitudes to sexuality cannot entirely explain observed differences between the sexes in sexual activity. Women were less likely to be married or in other intimate relationships than men, as reported by others.[10] As in other studies,[8 10 12] the proportion of elderly people reporting sexual activity was higher among married participants, especially in women.

In 2000 around two thirds of adults reported high sexual satisfaction, a substantial increase from 1976, especially in women. The proportion of women who reported no or low satisfaction decreased, whereas the proportion of men with low satisfaction increased. It could be speculated that it has become more permissible for men to admit failure in sexual matters or that a real difference exists.

It was beyond the scope of this study to examine in detail the reasons for self reported secular changes in sexual behaviour over 30 years. It could be speculated that the changes reflect higher educational levels and better socioeconomic status in the later birth cohorts. Furthermore, cohabiting and living apart became more common. The proportion never married decreased and the proportion divorced increased in the later born samples. These samples also experienced better general health.[17] When several of these factors were taken into consideration in a logistic regression analysis, birth year was still related to sexual activity, suggesting that several unidentified factors might be important. Changes in legislation may have influenced public attitudes to sexuality during the 20th century. Sex education became compulsory in schools in Sweden in 1955. By the end of the 1950s condoms were available in vending machines in public places. The "sexual revolution" followed in the 1960s, with the contraceptive pill and intrauterine devices.[18]

Strengths and limitations

This study was based on four population samples examined similarly over a 30 year period. The interviews were part of a comprehensive investigation on ageing and thus people were not recruited explicitly to talk about their sexuality.

This study had several limitations. Firstly, although higher than in most studies on sexual behaviour, the response rate declined from 80% to 65% during the 30 year period. No differences were identified between responders and non-responders for several factors. The secular trends in reported sexual behaviour over the study period were, however, pronounced. We cannot exclude the possibility that non-responders had more sexual problems than responders. Secondly, we examined 70 year olds thus we cannot draw conclusions on sexual behaviour before this age. Thirdly, sexual behaviour is a sensitive matter to report. Semistructured interviews were, however, done by doctors or psychiatric research nurses. Fourthly, the first three cohorts were examined by psychiatrists and the fourth by nurses, but differences were not large. Fifthly, changes in evaluations of responses over time may have influenced the results. One researcher (IS) was trained by those who carried out the examinations in the 1970s, who in turn trained those doing the examinations in 1992 and 2000. Inter-rater reliability between the researcher and examiners in the 1970s and 1990s was high. Sixthly, the study is based on self report, lending itself to reporting bias. More positive attitudes to sexuality in 70 year olds in later born cohorts might have resulted in more participants reporting intercourse. It is possible that our results reflect a more open minded attitude in society to sexual matters rather than real changes in sexual behaviour. Seventhly, the definition of sexual activity was limited to intercourse between heterosexuals. Thus we cannot generalise our results to other types of sexuality. As we aimed to describe secular trends, we were limited to those questions used in the 1970s. Finally, depression is common in elderly people and is well known to affect sexual activity. Our results for prevalence, however, did not change when we excluded depressed participants, and year of birth was still related to sexual activity in 70 year olds when depression was controlled for in logistic regression analyses including all four samples.

Conclusions

Self reported quantity and quality of sexual experiences among 70 year olds improved over a 30 year period. At the same time, a relatively large proportion of participants had ceased having intercourse. Our study, however, shows that most elderly people

consider sexual activity and associated feelings a natural part of later life.

Contributors: See bmj.com.
Funding: This study was supported by grants from the Swedish Council for Working Life and Social Research (No 2001-2835, 2001-2646, 2003-0234, 2004-0150, 2004-0145, 2006-0596, and 2006-0020), the Alzheimer's Association Stephanie B Overstreet Scholars (IIRG-00-2159), the Swedish Research Council (No 11267, 2005-8460, and 825-2007-7462), the Bank of Sweden Tercentary Foundation, Stiftelsen för Gamla Tjänarinnor, and Handlanden Hjalmar Svenssons Forskningsfond. The sponsors had no role in the study design, data collection, data analyses, the interpretation of data, the writing of the report, or the decision to submit the article for publication.
Competing interests: None declared.
Ethical approval: This study was approved by the ethics committee for medical research at Gothenburg University.
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Persson G. Sexuality in a 70-year-old urban population. *J Psychosom Res* 1980;24:335-42.
2 Nilsson LV. Prevalence of mental disorder in a 70-year-old urban sample. A cohort comparison. *J Clin Exp Gerontol* 1983;5:101-20.
3 Rinder L, Roupe S, Steen B, Svanborg A. Seventy-year-old people in Gothenburg. A population study in an industrialized Swedish city. *Acta Med Scand* 1975;198:397-407.
4 Turner CF, Danella RD, Rogers SM. Sexual behavior in the United States 1930-1990: trends and methodological problems. *Sex Transm Dis* 1995;22:173-90.
5 Lewin B, Fugl-Meyer K, Helmius G, Lalos A, Månsson S, the National Institute of Public Health. *Sexuality in Sweden: about sex life in Sweden 1996*. 3rd ed. Stockholm: National Institute of Public Health, 1998. (Swedish language.)
6 Caron SL, Moskey EG. Changes over time in teenage sexual relationships: comparing the high school class of 1950, 1975, and 2000. *Adolescence* 2002;37:515-26.
7 George LK, Weiler SJ. Sexuality in middle and late life. The effects of age, cohort, and gender. *Arch Gen Psychiatry* 1981;38:919-23.
8 Diokno AC, Brown MB, Herzog AR. Sexual function in the elderly. *Arch Intern Med* 1990;150:197-200.
9 Nicolosi A, Buvat J, Glasser DB, Hartmann U, Laumann EO, Gingell C. Sexual behaviour, sexual dysfunctions and related help seeking patterns in middle-aged and elderly Europeans: the global study of sexual attitudes and behaviors. *World J Urol* 2006;24:423-8.
10 Lindau ST, Schumm LP, Laumann EO, Levinson W, O'Muircheartaigh CA, Waite LJ. A study of sexuality and health among older adults in the United States. *N Engl J Med* 2007;357:762-74.
11 Laumann EO, Nicolosi A, Glasser DB, Paik A, Gingell C, Moreira E, et al. Sexual problems among women and men aged 40-80 y: prevalence and correlates identified in the global study of sexual attitudes and behaviors. *Int J Impot Res* 2005;17:39-57.
12 Verwoerdt A, Pfeiffer E, Wang HS. Sexual behavior in senescence. Changes in sexual activity and interest of aging men and women. *J Geriatr Psychiatry* 1969;2:163-80.
13 Verwoerdt A, Pfeiffer E, Wang HS. Sexual behavior in senescence. II. Patterns of sexual activity and interest. *Geriatrics* 1969;24:137-54.
14 Pfeiffer E, Davis GC. Determinants of sexual behavior in middle and old age. *J Am Geriatr Soc* 1972;20:151-8.
15 Kivelä S, Pahkala K, Honkakoski A. Sexual desire, intercourse, and related factors among elderly Finns. *Nordisk sexologi* 1986:18-27.
16 Pfeiffer E, Verwoerdt A, Wang HS. The natural history of sexual behavior in a biologically advantaged group of aged individuals. *J Gerontol* 1969;24:193-8.
17 Wilhelmson K, Allebeck P, Steen B. Improved health among 70 year olds: comparison of health indicators in three different birth cohorts. *Aging Clin Exp Res* 2002;14:361-70.
18 Sundström K, the National Institute of Public Health. *How do others do? About sexuality and relationship in the 1990s*. Stockholm: National Institute of Public Health, 2000. (Swedish language.)

Accepted: 12 May 2008

Prognosis in patients with recent onset low back pain in Australian primary care: inception cohort study

Nicholas Henschke,[1,2] Christopher G Maher,[1,2] Kathryn M Refshauge,[2] Robert D Herbert,[1,2] Robert G Cumming,[3] Jane Bleasel,[4] John York,[4] Anurina Das,[2] James H McAuley[1,2]

[1]Musculoskeletal Division, The George Institute for International Health, Sydney, Australia
[2]Back Pain Research Group, Faculty of Health Sciences, University of Sydney
[3]School of Public Health, University of Sydney
[4]Department of Rheumatology, Royal Prince Alfred Hospital, Sydney

Correspondence to: C G Maher, PO Box M201, Camperdown NSW 2050, Australia
cmaher@george.org.au

Cite this as: *BMJ* 2008;337:a171
doi:10.1136/bmj.a171

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a171

ABSTRACT

Objective To estimate the one year prognosis and identify prognostic factors in cases of recent onset low back pain managed in primary care.
Design Cohort study with one year follow-up.
Setting Primary care clinics in Sydney, Australia.
Participants An inception cohort of 973 consecutive primary care patients (mean age 43.3, 54.8% men) with non-specific low back pain of less than two weeks' duration recruited from the clinics of 170 general practitioners, physiotherapists, and chiropractors.
Main outcome measures Participants completed a baseline questionnaire and were contacted six weeks, three months, and 12 months after the initial consultation. Recovery was assessed in terms of return to work, return to function, and resolution of pain. The association between potential prognostic factors and time to recovery was modelled with Cox regression.
Results The follow-up rate over the 12 months was more than 97%. Half of those who reduced their work status at baseline had returned to previous work status within 14 days (95% confidence interval 11 to 17 days) and 83% had returned to previous work status by three months. Disability (median recovery time 31 days, 25 to 37 days) and pain (median 58 days, 52 to 63 days) took much longer to resolve. Only 72% of participants had completely recovered 12 months after the baseline consultation. Older age, compensation cases, higher pain intensity, longer duration of low back pain before consultation, more days of reduced activity because of lower back pain before consultation, feelings of depression, and a perceived risk of persistence were each associated with a longer time to recovery.
Conclusions In this cohort of patients with acute low back pain in primary care, prognosis was not as favourable as claimed in clinical practice guidelines. Recovery was slow for most patients. Nearly a third of patients did not recover from the presenting episode within a year.

INTRODUCTION

There is evidence that the type of advice given to patients can alter the course of an episode of low back pain. For this reason, most management guidelines recommend that patients should be reassured that they have a favourable prognosis. Several guidelines state that 90% of patients recover within six weeks.[1 2] Such

statements, however, might be too optimistic. While patients typically improve rapidly, the risk of developing chronic low back pain (that is, pain persisting for more than three months) is uncertain. Estimates of this risk vary from 2%[3] to 56%.[4]

To provide individualised advice, it is necessary to consider prognostic factors. All guidelines for low back pain recommend identification of adverse prognostic factors, commonly described as "yellow flags."

The lack of consensus regarding the prognosis and prognostic factors for recent onset low back pain has been attributed to methodological shortcomings of previous studies.[5-7]

We conducted a cohort study with the primary aim of determining the long term (one year) prognosis for people with recent onset low back pain presenting to primary care clinicians. Our secondary aim was to identify patients' characteristics that could be readily assessed by a primary care clinician and were associated with poor prognosis.

METHODS

The protocol for this study has been published previously.[8] We recruited an inception cohort of 973 participants from a socioeconomically diverse region in the Sydney metropolitan area of Australia. We invited all general practitioners, physiotherapists, and chiropractors within the study area to participate. Participating clinicians screened all patients with the primary complaint of low back pain who presented to their clinics from November 2003 to July 2005.

Low back pain was defined as pain in the area bounded superiorly by T12 and inferiorly by the buttock crease, lasting for more than 24 hours but less than two weeks, and preceded by a period of at least one month without back pain. See bmj.com for exclusions. The clinicians were given a copy of the most recent clinical guidelines for low back pain and asked to follow the guidelines when appropriate.

Baseline data were collected at the first consultation with the primary care clinician. These data were used to describe the cohort and to evaluate putative predictors of outcome. The individual variables were grouped into seven factors. Low back pain and disability were also measured at baseline with adaptations of items 7 and 8 of the SF-36. Clinicians were paid for participating.

Researchers conducted follow-up assessments by telephone at six weeks, three months, and 12 months after the initial assessment. We sampled three dimensions of recovery—pain intensity, disability, and work status—which participants were asked to rate at each time point. Interviewers also established whether the patient had recovered on each of these dimensions and if so the date of recovery. A fourth measure of recovery —"complete recovery"—required the patient to recover on all three dimensions.

Data analysis—We used the dates on which participants returned to pre-injury work status and/or had no disability and/or had no pain to construct survival curves. Median survival time (days to recovery) was determined for each of the three recovery measures individually and for attainment of all three recovery

Cox regression model for time to complete recovery from acute low back pain (LBP) with hazard ratios (HR) and 95% confidence intervals

	Crude (unadjusted)		Adjusted	
Variable	HR (95% CI)	P value	HR (95% CI)	P value
Age (years)	1.00 (0.99 to 1.00)	0.090	0.99 (0.99 to 1.00)	0.004
Male	1.05 (0.91 to 1.23)	0.500	1.01 (0.86 to 1.18)	0.900
Pain/disability (χ^2=31.32, P<0.001)				
Pain intensity*	0.79 (0.73 to 0.86)	<0.001	0.86 (0.77 to 0.96)	0.009
Interference with function†	0.85 (0.90 to 0.91)	<0.001	0.96 (0.88 to 1.05)	0.339
Psychological (χ^2=81.51, P<0.001)				
Pain control§	1.04 (1.01 to 1.07)	0.010	1.02 (0.99 to 1.05)	0.267
Tension/anxiety¶	0.94 (0.91 to 0.96)	<0.001	1.02 (0.99 to 1.06)	0.208
Feelings of depression¶	0.91 (0.89 to 0.93)	<0.001	0.94 (0.91 to 0.97)	<0.001
Risk of persistence¶	0.89 (0.87 to 0.92)	<0.001	0.92 (0.89 to 0.95)	<0.001
Current history (χ^2=36.72, P<0.001)				
Compensable LBP¶	0.56 (0.45 to 0.69)	<0.001	0.59 (0.47 to 0.74)	<0.001
Currently taking medication for LBP	0.75 (0.65 to 0.88)	<0.001	0.96 (0.81 to 1.14)	0.657
Days of reduced activity due to LBP¶	0.96 (0.93 to 0.99)	0.005	1.04 (1.00 to 1.08)	0.033
Leg pain¶	0.71 (0.59 to 0.86)	0.001	0.90 (0.70 to 1.16)	0.408
No of pain sites‡	0.83 (0.75 to 0.91)	<0.001	0.92 (0.81 to 1.03)	0.147
Duration of episode	0.97 (0.95 to 0.99)	0.030	0.97 (0.94 to 1.0)	0.033
Clinical red flags (not included in model)				
No of positive red flags	0.95 (0.91 to 1.00)	0.048	—	

*Pain intensity scale: 1=none, 2=very mild, 3=mild, 4=moderate, 5=severe, 6=very severe.
†Disability scale: 1=not at all, 2=little bit, 3=moderate, 4=quite a bit, 5=extreme.
§Rated on scale from 0-10, with higher score indicating better ability to control pain.
¶Rated on scales from 0-10, with higher scores indicating more tension and anxiety, more feelings of depression, or higher risk of persistent pain.
‡One point for each pain site: neck, shoulder, upper back, lower back, and leg.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Clinical practice guidelines suggest that recovery from an episode of recent onset low back pain is usually rapid and complete

Recent systematic reviews suggest that the risk of developing chronic low back pain is uncertain

WHAT THIS STUDY ADDS

In this cohort of patients, recovery from recent onset low back pain was much slower than has been reported and nearly a third did not recover within a year

Older age, back pain associated with compensation cases, higher pain intensity, longer duration of low back pain before consultation, more days of reduced activity because of low back pain before consultation, feelings of depression, and a perceived risk of persistence were all associated with poorer prognosis

measures. We used Cox regression to identify independent associations between the seven factors and the time to complete recovery from acute low back pain. See bmj.com for further details.

RESULTS

A total of 1254 clinicians in the study region were identified and contacted. Of these, 170 (73 general medical practitioners, 77 physiotherapists, and 20 chiropractors) were trained in the study protocol. They screened 3184 consecutive patients with low back pain from November 2003 to July 2005. Of these, 973 patients (mean age 43.3, 54.8%) were eligible to participate. The follow-up rate remained above 97% over the 12 month period. Participants for whom we did not have complete follow-up data and who had not reported recovery from their low back pain were censored at the time of their last follow-up.

Full baseline demographic and clinical features of participants are on bmj.com. We used the participant's postcode and data from the Australian census to judge socioeconomic disadvantage: 21.4% fell in the most disadvantaged quarter, 8.4% in the second quarter, 14.8% in the third quarter, and 55.1% in the fourth quarter.

There were 770 (79.5%) participants who reported working before the onset of their episode of acute low back pain. Of these, 291 (37.8% of workers, 30.0% of the total cohort) reported changing their work status as a result of their low back pain. The median time to return to previous work hours and duties for these 291 participants was 14 days (95% confidence interval 11 to 17 days). The cumulative probability of returning to pre-back pain work hours and duties for those who reduced their work status at baseline because of low back pain was 74.6% at six weeks, 83.2% at 12 weeks, and 89.5% one year after consulting a primary care clinician for acute low back pain.

The median time to recovery in terms of disability was 31 days (25 to 37 days). By six weeks the cumulative probability of having no disability was 54.9%. This probability increased to 73.3% by 12 weeks, and 83.3% by one year. The median time to recovery in terms of pain after an episode of acute low back pain was 58 days (53 to 63 days). The cumulative probability of being pain-free was 39.9% by six weeks, 58.2% by 12 weeks, and 72.5% by one year.

Complete recovery from recent onset low back pain, determined by recovery on all three dimensions (return to work, no disability, and no pain) took a median time of 59 days (53 to 65 days). Six weeks after presentation to primary care, the cumulative probability of recovery was 39.0%. By 12 weeks the probability was 57.4%, and this increased to 71.8% by one year.

At six weeks, three months, and 12 months, 40%, 52%, and 57% of participants reported being pain-free; 60%, 71%, and 75% reported being disability-free. Immediately before the onset of the episode 77% were working full time and this reduced to 48% at baseline, rising to 69%, 72%, and 72% at six weeks, three months, and 12 months. At 12 months the participants' responses to the question "If you had to live with the symptoms you have right now, how would you feel about it?" were generally positive, though 133 of the 969 participants reported feeling very dissatisfied and 106 somewhat dissatisfied.

After adjustment for age, sex, intensity of pain, and interference with function, psychological characteristics were most closely associated with time to recovery (table). Of the other factors, only factors related to current history further contributed significantly to the model. Seven individual variables were independently associated with time to recovery: age, intensity of pain, feelings of depression, risk of persistence, low back pain in compensation cases, days of reduced activity, and duration of the episode.

DISCUSSION

In this study of 12 month prognosis in patients with recent onset low back pain, recovery was typically much slower than previously reported. Nearly a third of patients did not recover from the presenting episode within a year. Return to work and recovery from disability and pain did not occur synchronously. We identified seven factors that were associated with speed of recovery and can be considered by clinicians when advising their patients about the prognosis for their episode of acute low back pain.

Strengths and weaknesses

We enrolled an inception cohort from the three main primary care providers who manage low back pain and measured pain, disability, and work status over a 12 month period with high rates of follow-up. Our previous review of prognostic studies of low back pain found that few studies of acute low back pain have achieved these benchmarks.[6] Socioeconomically disadvantaged people were under-represented in the cohort. Also we did not record participants' occupation so we were unable to assess whether this factor influenced the speed with which people returned to work.

Comparison with other research

There are only a few methodologically sound prognosis studies that have followed patients beyond three

months.[6] A Danish study also found that recovery was slow and incomplete.[4] In contrast, a French study reported that recovery was rapid.[3] We are unable to explain the marked difference in results. There are also difficulties comparing prognostic factors between these studies. None the less, all three studies report that compensation cases and high disability at baseline were adverse prognostic factors, and our study, and the French study, report that a previous episode of low back pain was an adverse prognostic factor. The Danish study reported that perceived risk of persistence was an adverse prognostic factor,[4] but in that study the clinician judged risk of persistence whereas in our study this judgment was made by the patient.

Implications for the guidelines

Our findings support the recommendations in clinical practice guidelines that clinicians should screen for adverse prognostic factors (yellow flags). Recovery did not occur synchronously in the three dimensions of return to work, interference with function, and pain status. Pain took the longest to resolve and the survival curves for recovery from pain and complete recovery were similar.

There has been little consensus regarding predictors of outcome from acute low back pain.[7 9] Prognostic information can be used to provide patient specific estimates of prognosis to individual patients in primary care.

Contributors: See bmj.com.
Funding: National Health and Medical Research Council of Australia.
Competing interests: None declared.
Ethical approval: University of Sydney human research ethics committee.
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Van Tulder M, Becker A, Bekkering T, Breen A, del Real MT, Hutchinson A, et al. Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. *Eur Spine J* 2006;15:S169-91.
2 Waddell G, Feder G, McIntosh A, Lewis M, Hutchinson A. *Low back pain evidence review*. 1st ed. London: Royal College of General Practitioners, 1996.
3 Coste J, Delecoeuillerie G, Cohen de Lara A, Le Parc JM, Paolaggi JB. Clinical course and prognostic factors in acute low back pain: an inception cohort study in primary care practice. *BMJ* 1994;308:577-80.
4 Schiottz-Christensen B, Nielsen GL, Hansen VK, Schodt T, Sorensen HT, Olesen F. Long-term prognosis of acute low back pain in patients seen in general practice: a 1-year prospective follow-up study. *Fam Pract* 1999;16:223-32.
5 Hestbaek L, Leboeuf-Yde C, Manniche C. Low back pain: what is the long-term course? A review of studies of general patient populations. *Eur Spine J* 2003;12:149-65.
6 Pengel L, Herbert R, Maher C, Refshauge K. Acute low back pain: systematic review of its prognosis. *BMJ* 2003;327:323-7.
7 McIntosh G, Frank J, Hogg-Johnson S, Hall H, Bombardier C. Low back pain prognosis: structured review of the literature. *J Occ Rehabil* 2000;10:101-15.
8 Henschke N, Maher CG, Refshauge KM, Herbert RD, Cumming RG, Bleasel J, et al. Prognosis of acute low back pain: design of a prospective inception cohort study. *BMC Musculoskelet Disord* 2006;7:54.
9 Kent PM, Keating JL. Can we predict poor recovery from recent-onset nonspecific low back pain? A systematic review. *Man Ther* 2008;13:12-28.

Accepted: 25 April 2008

Routine care of peripheral intravenous catheters versus clinically indicated replacement: randomised controlled trial

Joan Webster,[1,2] Samantha Clarke,[1] Dana Paterson,[1] Anne Hutton,[1] Stacey van Dyk,[1] Catherine Gale,[1] Tracey Hopkins[1]

EDITORIAL by Maki

[1]Centre for Clinical Nursing, Royal Brisbane and Women's Hospital, Herston, Qld 4029, Australia
[2]Research Centre for Clinical and Community Practice Innovation Griffith University, Centre for Clinical Nursing, Royal Brisbane and Women's Hospital
Correspondence to: J Webster
joan_webster@health.qld.gov.au

Cite this as: *BMJ* 2008;337:a339
doi:10.1136/bmj.a339

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a339.

ABSTRACT

Objective To compare routine replacement of intravenous peripheral catheters with replacement only when clinically indicated.
Design Randomised controlled trial.
Setting Tertiary hospital in Australia.
Participants 755 medical and surgical patients: 379 allocated to catheter replacement only when clinically indicated and 376 allocated to routine care of catheter (control group).
Main outcome measure A composite measure of catheter failure resulting from phlebitis or infiltration.
Results Catheters were removed because of phlebitis or infiltration from 123 of 376 (33%) patients in the control group compared with 143 of 379 (38%) patients in the intervention group; the difference was not significant (relative risk 1.15, 95% confidence interval 0.95 to 1.40). When the analysis was based on failure per 1000 device days (number of failures divided by number of days catheterised, divided by 1000), no difference could be detected between the groups (relative risk 0.98, 0.78 to 1.24). Infusion related costs were higher in the control group (mean $A41.02;£19.71;€24.80;$38.55) than intervention group ($A36.40). The rate of phlebitis in both groups was low (4% in intervention group, 3% in control group).
Conclusion Replacing peripheral intravenous catheters when clinically indicated has no effect the incidence of failure, based on a composite measure of phlebitis or infiltration. Larger trials are needed to test this finding using phlebitis alone as a more clinically meaningful outcome.
Registration number Australian New Zealand Clinical Trials Registry ACTRN12605000147684.

INTRODUCTION

Intravenous catheterisation is the most common invasive procedure among inpatients, with about half receiving intravenous therapy.[1] Between 2.3% and 67% of patients will develop phlebitis.[2-8] Despite the ubiquitous use of catheters and the almost universal acceptance of the need for routine replacement, the

practice has received little rigorous evaluation. We therefore carried out a randomised trial to compare routine changes of catheters with clinically indicated changes.

METHODS

We recruited inpatients from a general teaching hospital. Patients were eligible for inclusion if they were at least 18 years old, had no current bacteraemia, were not receiving immunosuppressive therapy, and were scheduled or expected to have a peripheral venous catheter for at least four days. The patient was the unit of measurement, therefore we entered those requiring multiple or consecutive catheters into the study once only.

We randomised patients either to routine care (control group), with catheters scheduled to be replaced every three days according to hospital policy, or to replacement only when clinically indicated (clinically indicated group). For each group we recorded reasons for replacement and deviations from the replacement protocol.

Participants were randomised by an investigator with no clinical involvement in the trial. We stratified by admission for an oncology related procedure. A research nurse telephoned a contact who was independent of the recruitment process for allocation consignment. Clinical staff were subsequently aware of the treatment group, to ensure that catheters were changed as scheduled and that those in the intervention group were not removed if functional and there was no sign of inflammation or infection.

The research nurse collected baseline personal, clinical, and catheter related data (see bmj.com). Choice of catheter and gauge was at the discretion of the professional inserting the catheter. Insertion sites were inspected daily by a nurse from the intravenous service and by ward nurses when solutions were changed or drugs added. To optimise the standardisation of reporting data, the nurse removing the catheter recorded the reasons for catheter removal on a specially designed form.

Any deviations from the protocol for catheter replacement were recorded. We collected data for up to five consecutive catheters for each patient.

Outcome measures

For the primary outcome we used a composite measure of catheter failure as a result of phlebitis or infiltration (see definitions on bmj.com).

Secondary outcomes included infusion related cost —those associated with catheters inserted for intermittent administration of intravenous drugs or for continuous infusion. For patients receiving intermittent drugs we calculated a total cost of $A16.4 per insertion (see bmj.com) and for those receiving a continuous infusion we calculated $A28.84 per insertion (see bmj.com). We also included other reasons for catheter failure as secondary outcomes: occlusion or blockage, local infection at the insertion site, or catheter related bloodstream infection.[9] We also included phlebitis and infiltration as independent secondary outcomes.

Statistical analysis

Trial data were analysed on an intention to treat basis. We calculated the relative risks (95% confidence intervals) for the proportion of patients with a failed catheter and for the individual factors that made up the composite measure. We also calculated the failure rate for each group per 1000 device days (number of failures divided by number of days catheterised, divided by 1000), which is a more meaningful measure for this outcome. When appropriate, we used a two sided Fisher's exact test to compare discrete data; results are presented as P values. We used the independent sample *t* test to compare the differences in the infusion related costs and total catherisation time between groups.

RESULTS

Overall, 755 of 1620 potentially eligible participants (46.6%) were included in the trial (see bmj.com): 376 were randomised to routine care of catheters (control group) and 379 to replacement of catheters only when clinically indicated (intervention group). Twenty two participants (6%) in the intervention group had catheters changed routinely. One hundred and twenty five participants (33%) in the control group had a catheter in place for more than 72 hours. Follow-up from medical records was possible for all participants.

The groups were similar for baseline personal, clinical, and catheter related characteristics for most risk factors (see bmj.com). The intervention group had higher rates for a history of phlebitis and presence of a wound infection or infected ulcer.

Each catheter was in place for a greater mean length of time in the intervention group than in the control group (see bmj.com). As a result more catheters were placed in the control group (n=749) than in the intervention group (n=679), despite the average number of intravenous therapy days being less in the control group: 6.3 (SD 5.1) in the intervention group versus 5.4 (SD 3.8) in the control group. The number of days catheterised was 2020 in the control group and 2393 in the intervention group.

Effect of intervention

Overall, 123 (33%) participants in the control group and 143 (38%) in the intervention group had catheters removed because of phlebitis or infiltration (table); the difference was not significant (relative risk 1.15, 95% confidence interval 0.95 to 1.40). When the analysis was based on failure per 1000 device hours, no difference could be detected between the groups. Infusion related costs per episode of care were higher in the control group than in the intervention group (mean $A41.02 *v* mean $A36.40). Both groups had low rates of phlebitis (4% in intervention group, 3% in

Outcomes for patients allocated to routine care of intravenous catheter (control group) or to replacement of catheter only when clinically indicated (intervention group). Values are numbers (percentages) of patients unless stated otherwise

Outcomes	Intervention group (n=379)	Control group (n=376)	Relative risk (95% CI)
Primary:			
Catheter failure per person	143 (38)	123 (33)	1.15 (0.95 to 1.40)
Catheter failure per 1000 device days	59.8	60.9	
Secondary:			
Mean (SD) intravenous cost per catheter	41.05 (26.6)	46.22 (28.7)	−5.16* (−9.12 to −1.21)
Phlebitis	16 (4)	12 (3)	1.32 (0.63 to 2.76)
Infiltration	135 (36)	120 (32)	1.12 (0.91 to 1.36)
Blockage	30 (8)	20 (5)	1.49 (0.86 to 2.57)
Local infection	2 (1)	0	4.96 (0.24 to 102.98)
Suspected bloodstream infection	1 (0.3)	1 (0.3)	0.99 (0.06 to 15.80)

*Mean difference.

control group). Infiltration was the most common reason for failure (36% in intervention group, 32% in control group). A total of 196 (26%) catheters in control participants were replaced after three days, according to hospital policy, despite functioning well.

DISCUSSION

The routine replacement of peripheral intravenous catheters has no effect on the incidence of catheter failure, on the basis of a composite measure of phlebitis or infiltration. The result replicates findings from an earlier study by us, which used narrower inclusion criteria but a broader definition of failure.[10] These studies have increased our confidence in changing intravenous lines according to signs and symptoms, rather than using predetermined times.[11] Changing our policy would bring the practice in adults in line with recommendations from the Centers for Disease Control and Prevention for changing peripheral intravenous lines in children—that is, replace catheters only when clinically indicated.

Our overall combined rate for phlebitis and infiltration was 35%, similar to other reports.[12 13] The phlebitis rate in both groups was on the low side of ranges reported in recent studies,[14 15] despite our population being elderly and almost 75% having at least one comorbidity. Reported rates depend on definitions used, and although we applied a standard definition, interpretation of signs and symptoms could be affected by subjectivity or omission of reporting. It is perhaps more useful to use the composite measure of infiltration or phlebitis to avoid misdiagnosis.

Despite group allocation, participants showed little difference in dwell times. Two factors contribute to this. Firstly, it is not possible to modify all routinely scheduled changes precisely 72 hours after insertion. Secondly, many of the catheters in the intervention group failed before 72 hours—although catheters remained in place longer than in the control group, the average dwell time was within the 72-96 hours recommended by the Centers for Disease Control and Prevention. This confirms that all catheters fail eventually but that many remain functional for prolonged periods (about 3% remained trouble free for over seven days and some for as long as two weeks). We therefore believe that routinely changing catheters may be an unnecessary, painful, and costly intervention.

Potential cost savings of about 25% for infusion related costs could be made if our policy was to be changed in line with recent evidence. Cost estimates used in our study were conservative, based on a simple intravenous event.

The study was not sufficiently powered to show differences in our secondary clinical outcomes. Despite this, non-significant results favoured the control group for lower rates of phlebitis, blockage, and local infection.

Strengths and limitations

The major strengths of the study were the processes used to eliminate selection bias, ensure allocation concealment, and ensure that the study was adequately powered to detect differences in our primary outcome. We also included a range of participants and did not impose caveats on how or by whom catheters should be inserted. This was to match normal practice and to enable the extrapolation of results to other inpatient populations. We enrolled 47% of eligible patients

WHAT IS ALREADY KNOWN ON THIS TOPIC

Peripheral intravenous catheterisation is the most common invasive procedure among inpatients

Changing catheters every three days to prevent infection is standard procedure but the practice has not been rigorously tested

WHAT THIS STUDY ADDS

Catheters may be safely left in place for longer than 72 hours if no contraindications are present

When catheters are replaced only when clinically indicated 25% of infusion related costs are saved

compared with about 25% in the earlier trial, with no losses to follow-up.

The study would have been strengthened if monitoring of outcomes had been more stringent. As it was, we extracted most of the outcome data from medical records. A more standardised approach would have been preferable, using staff trained in the process and data collected in real time. Also, outcome assessment was done by people who were not blinded to group allocation. Although catheters were removed by ward or intravenous service staff, part of their normal practice is to record reasons for removal in the patient's medical record. To falsify records because of group allocation would be unlikely. Finally, the study was not powered to study differences in secondary outcomes. Phlebitis alone would have been a more clinically important end point but we were limited by funding.

Contributors: See bmj.com.
Funding: Queensland Nursing Council research grant (RAN:0513) and the RBWH Research Foundation.
Competing interests: None declared.
Ethical approval: Royal Brisbane and Women's Hospital human research ethics committee (No 2005/93).
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Bregenzer T, Conen D, Sakmann P, Widmer AF. Is routine replacement of peripheral intravenous catheters necessary? *Arch Intern Med* 1998;158:151-6.
2 Tripepi-Bova KA, Woods KD, Loach MC. A comparison of transparent polyurethane and dry gauze dressings for peripheral IV catheter sites: rates of phlebitis, infiltration, and dislodgment by patients. *Am J Crit Care* 1997;6:377-81.
3 Chee S, Tan W. Reducing infusion phlebitis in Singapore hospitals using extended life end-line filters. *J Infus Nurs* 2002;25:95-104.
4 Tager IB, Ginsberg MB, Ellis SE, Walsh NE, Dupont I, Simchen E, et al. An epidemiologic study of the risks associated with peripheral intravenous catheters. *Am J Epidemiol* 1983;118:839-51.
5 Martinez JA, Fernandez P, Rodriguez E, Sobrino J, Torres M, Nubiola A, et al. Intravenous cannulae: complications arising from their use and analysis of their predisposing factors. *Med Clin (Barc)* 1994;103:89-93.
6 Lipsky BA, Peugeot RL, Boyko EJ, Kent DL. A prospective study of Staphylococcus aureus nasal colonization and intravenous therapy-related phlebitis. *Arch Intern Med* 1992;152:2109-12.
7 White SA. Peripheral intravenous therapy-related phlebitis rates in an adult population. *J Intraven Nurs* 2001;24:19-24.
8 Shimandle RB, Johnson D, Baker M, Stotland N, Karrison T, Arnow PM. Safety of peripheral intravenous catheters in children. *Infect Control Hosp Epidemiol* 1999;20:736-40.
9 Cornely OA, Bethe U, Pauls R, Waldschmidt D. Peripheral Teflon catheters: factors determining incidence of phlebitis and duration of cannulation. *Infect Control Hosp Epidemiol* 2002;23:249-53.
10 Webster J, Lloyd S, Hopkins T, Osborne S, Yaxley M. Developing a research base for intravenous peripheral cannula re-sites (DRIP trial). A randomised controlled trial of hospital in-patients. *Int J Nurs Stud* 2007;44:664-71.
11 O'Grady NP, Alexander M, Dellinger EP, Gerberding JL, Heard SO, Maki DG, et al. Guidelines for the prevention of intravascular catheter-related infections. *Infect Control Hosp Epidemiol* 2002;23:759-69.
12 Wright A, Hecker J. Infusion failure caused by phlebitis and extravasation. *Clin Pharm* 1991;10:630-4.
13 Garland JS, Dunne WM Jr, Havens P, Hintermeyer M, Bozzette MA, Wincek J, et al. Peripheral intravenous catheter complications in critically ill children: a prospective study. *Pediatrics* 1992;89:1145-50.
14 Malach T, Jerassy Z, Rudensky B, Schlesinger Y, Broide E, Olsha O, et al. Prospective surveillance of phlebitis associated with peripheral intravenous catheters. *Am J Infect Control* 2006;34:308-12.
15 Nassaji-Zavareh M, Ghorbani R. Peripheral intravenous catheter-related phlebitis and related risk factors. *Singapore Med J* 2007;48:733-6.

Accepted: 7 May 2008

Bullous pemphigoid and pemphigus vulgaris—incidence and mortality in the UK: population based cohort study

S M Langan,[1] L Smeeth,[2] R Hubbard,[3] K M Fleming,[3] C J P Smith,[3] J West[3]

EDITORIAL by Joly

[1]Centre of Evidence-based Dermatology, University of Nottingham, Queen's Medical Centre, Nottingham NG7 2UH
[2]Department of Epidemiology and Population Health, Centre for Population Studies, Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London
[3]Division of Epidemiology and Public Health, University of Nottingham
Correspondence to: S Langan
sinead.langan@nottingham.ac.uk

Cite this as: *BMJ* 2008;337:a180
doi:10.1136/bmj.a180

ABSTRACT

Objective To determine the incidence of and mortality from bullous pemphigoid and pemphigus vulgaris in the United Kingdom.
Design Retrospective historical cohort study.
Setting Computerised medical records from the health improvement network, a large population based UK general practice database.
Participants Patients with pemphigus vulgaris and bullous pemphigoid diagnostic codes and age, sex, and practice matched controls.
Main outcome measures Incidence and mortality compared with the control population by calendar period, age group, sex, geographical region, and degree of social deprivation.
Results 869 people with bullous pemphigoid and 138 people with pemphigus vulgaris were identified. The median age at presentation for bullous pemphigoid was 80 (range 23-102) years, and 534 (61%) patients were female. The median age at presentation for pemphigus vulgaris was 71 (21-102) years, and 91 (66%) patients were female. Incidences of bullous pemphigoid and pemphigus vulgaris were 4.3 (95% confidence interval 4.0 to 4.6) and 0.7 (0.6 to 0.8) per 100 000 person years. The incidence of bullous pemphigoid increased over time; the average yearly increase was 17% (incidence rate ratio=1.2, 95% confidence interval 1.1 to 1.2). An average yearly increase in incidence of pemphigus vulgaris of 11% (incidence rate ratio=1.1, 1.0 to 1.2) occurred. The risk of death for patients with bullous pemphigoid was twice as great as for controls (adjusted hazard ratio=2.3, 95% confidence interval 2.0 to 2.7). For pemphigus vulgaris, the risk of death was three times greater than for controls (adjusted hazard ratio=3.3, 2.2 to 5.2).
Conclusions Incidences of bullous pemphigoid and pemphigus vulgaris are increasing. The reasons for the changes in incidence are not clearly understood but have implications for identifying causative factors. Both disorders are associated with a high risk of death. Previous estimates may have underestimated the risk of death associated with these diseases.

Crude and adjusted survival analysis for patients with bullous pemphigoid and pemphigus vulgaris

	Deaths	Person years	Crude hazard ratio (95% CI)	Adjusted hazard ratio (95% CI)
Bullous pemphigoid				
Controls	604	9765	1.0	1.0
Cases	264	1993	2.11 (1.82 to 2.44)	2.29 (1.98 to 2.65)
Pemphigus vulgaris				
Controls	58	1763	1.0	1.0
Cases	36	380	2.82 (1.86 to 4.27)	3.38 (2.21 to 5.17)

INTRODUCTION

Previous studies of bullous pemphigoid have reported incidences between 0.2 and 3 per 100 000 person years; these are higher in older age groups, and some studies report associations with sex.[1-4] Wide variation in mortality is reported, with one year mortality varying between 6% in the United States and 41% in France.[5,6]

Hospital based studies of the epidemiology of pemphigus vulgaris report incidence varying between 0.076 and 1.6 per 100 000 person years.[7,8] Several studies have suggested a higher incidence in women.[8,9] Two studies report one year mortality of 4.8% and 54%, with no clear estimate of overall disease specific mortality.[10,11]

The published studies are mainly small and hospital based. We carried out a large population based study in people with bullous pemphigoid and pemphigus vulgaris to provide data on incidence and demographic data in the UK and to resolve controversies about mortality from these autoimmune disorders, which cause blistering of the skin and mucous membranes.

METHODS

The health improvement network is a computerised longitudinal general practice database with demographic data similar to the general population. The version we used contained data from 328 general practices that use "In Practice Vision" software. The quality of the data has been validated.[12]

Study population

Between 1996 and 2006, we identified all patients with a diagnosis of bullous pemphigoid or pemphigus vulgaris from the health improvement network database. In order to exclude prevalent cases, we imposed a lag period of three months after registration with their general practitioner to differentiate between incident and prevalent cases. We selected a control group of up to four controls per case, matched randomly by age, sex, and general practice, and contributing data to the database. We excluded people aged under 20 years. We assigned a date of "pseudodiagnosis" to controls, which was the date of diagnosis for the matched case.

Outcomes

We recorded incident diagnoses and dates of death. The follow-up period began on the date of first diagnosis or "pseudodiagnosis."

Statistical analysis

Incidence—We calculated incidences by age, categorised into 10 year age bands; sex; and calendar period. We used multivariate regression to model incidence rate ratios, adjusting for changes in age, sex structure, and calendar period over time. We applied calculated incidences for both diseases to the UK population totals for the years 2001-5 to estimate the number of new cases a year.

Mortality—We identified all deaths in the two case populations and matched controls and calculated one year mortality and five year survival rates. We used Cox regression to compare the mortality of cases and controls, adjusting for age, sex, and calendar period.

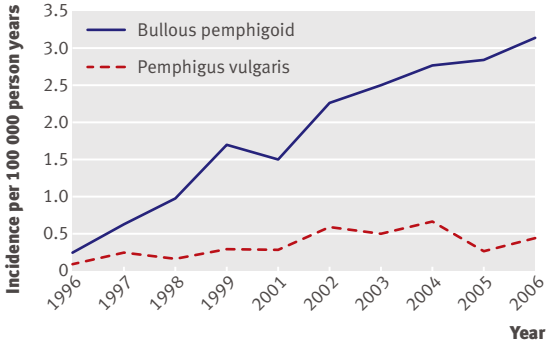
Office for National Statistics data—We calculated expected numbers of incident cases and deaths for both diseases by applying study rates to the Office for National Statistics population (2001-5) and comparing results with Office for National Statistics data.

RESULTS

Our cohort included 869 people with bullous pemphigoid contributing 1993 person years and 3453 matched controls contributing 9765 person years. The median age at first presentation for bullous pemphigoid was 80 (range 23-102) years, and 534 (61%) patients were women. One hundred and thirty eight people with pemphigus vulgaris contributed 380 person years, and 551 matched controls contributed 1763 person years. The median age at first presentation for pemphigus vulgaris was 71 (21-102) years, and 91 (66%) patients were women. The median length of follow-up for people with bullous pemphigoid was 1.6 (range 0-9) years, and that for pemphigus vulgaris was 2.0 (0-10.4) years. Losses to follow-up were similar in cases and controls in both diseases (by the end of year 2, bullous pemphigoid cases and controls 34% lost; pemphigus vulgaris cases 30% lost, controls 31% lost).

Bullous pemphigoid

Incidence—The crude incidence of bullous pemphigoid was 4.28 (95% confidence interval 4.01 to 4.58) per 100 000 person years. Incidence increased with age and in later calendar periods (figure). The increased



Age adjusted rates of bullous pemphigoid and pemphigus vulgaris, with direct standardisation to European standard population

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a180

WHAT IS ALREADY KNOWN ON THIS TOPIC

Little is known about the epidemiology of bullous pemphigoid and pemphigus vulgaris

Bullous pemphigoid is known to be associated with high mortality; mortality of pemphigus vulgaris in western Europe is not known

WHAT THIS STUDY ADDS

The incidences of bullous pemphigoid and pemphigus vulgaris are increasing, but the reasons for this are not clear

Mortality of bullous pemphigoid is twice that of the general population, and mortality of pemphigus vulgaris is three times as high as in the general population

Previous measures may have underestimated the burden of these diseases, in terms of both incidence and risk of death

incidence over calendar time persisted after adjustment for age group and sex (P for trend <0.0001). The estimated increase in the incidence of bullous pemphigoid per increase in calendar year was 17% (rate ratio 1.17, 95% confidence interval 1.14 to 1.20; P<0.0001) after adjustment for sex and age group, corresponding to a 4.8-fold increase in incidence over the 11 year period or a crude increase of 5.2-fold during the study. Applying this data to the UK population between 2001 and 2005 gave an average of 2996 (95% confidence interval 2485 to 3602) new cases a year.

Mortality—The table shows mortality for the bullous pemphigoid cohort. The overall crude mortality was 131.99 (95% confidence interval 116.96 to 148.95) per 1000 person years for bullous pemphigoid and 61.86 (57.12 to 66.99) per 1000 person years for controls. The absolute excess mortality in the bullous pemphigoid cohort was 70 per 1000 person years. The one year mortality for bullous pemphigoid was 19% (95% confidence interval 16.2% to 21.8%). The table shows crude and adjusted hazard ratios. We found no evidence of effect modification by calendar time.

Pemphigus vulgaris

Incidence—The crude incidence of pemphigus vulgaris was 0.68 (0.58 to 0.80) per 100 000 person years. Incidence was higher in women and in older age groups. We found evidence of an estimated 11% increase in incidence per calendar year (incidence rate ratio 1.11 (1.04 to 1.17) per year; P=0.001) (figure). We saw a slight reduction in incidence in the most recent period from 3.5 (1.9 to 6.2) in 2002-4 to 2.4 (1.2 to 4.5) in 2005-6. Applying these data to the UK population between 2001 and 2005 gave an average of 467 (297 to 643) new cases of pemphigus vulgaris a year.

Mortality—The table shows mortality for the pemphigus vulgaris cohort. The overall crude mortality was 94.64 (68.268 to 131.205) per 1000 person years for pemphigus vulgaris and 32.89 (25.43 to 42.55) per 1000 person years for controls. The absolute excess mortality in the pemphigus vulgaris cohort was 62 per 1000 person years. The one year mortality for pemphigus vulgaris was 12% (8% to 19%). The table shows crude and adjusted hazard ratios.

Comparison with Office for National Statistics mortality data

Applying the mortality data to the UK population data gave 1977 deaths in bullous pemphigoid cases and 221 deaths in pemphigus vulgaris cases (2001-5). These are much higher figures than the 190 and 36 deaths attributed to these diseases in the Office for National Statistics dataset.

DISCUSSION

We found a substantial increase in the incidence of both bullous pemphigoid and pemphigus vulgaris between 1996 and 2005. This increase was not associated with any reduction in mortality over the calendar periods and is therefore unlikely to be explained by ascertainment bias leading to the diagnosis of less severe cases. Our study therefore suggests that almost 3000 new cases of bullous pemphigoid and 500 of pemphigus vulgaris occur each year. Age and sex adjusted mortality was more than twice as high in people with bullous pemphigoid compared with the general population and three times as high in those with pemphigus vulgaris. The reported mortality for bullous pemphigoid increased from 63 to 88 per 1000 person years between 2001 and 2005, and that for pemphigus vulgaris increased from 13 to 19 per 1000 person years.

Strengths and weaknesses

This is a large, population based study, which allows robust estimation of incidence and mortality. The size of the dataset gives sufficient power to exclude chance as the basis for the findings. The use of routinely collected data means that we could not validate the diagnoses or assess severity of disease. These diagnoses are made in secondary care, and therefore likely to be accurate. Previous studies of other conditions have confirmed the validity of diagnoses in both the general practice research database and the health improvement network. Limiting the dataset to practices considered "up to standard" will have improved the quality. Some of the cases included may be prevalent rather than incident cases—if this is the case, the results may be affected by survival bias and we may have underestimated the mortality associated with these diseases.

Comparison with other studies

The mean age of patients with bullous pemphigoid in our study is similar to other published data. However, with the exception of one regional study, all studies were hospital based; therefore, although they may have had excellent validity, they may be less representative of the disease in the population. The mean age of patients with pemphigus vulgaris in our study was much higher than previous reports. However, no data on the epidemiology of pemphigus vulgaris in western Europe have been published and all of the studies are from hospital centres in the form of either cross sectional surveys or case series.

The incidence of bullous pemphigoid in our study is similar to that described by others[2]; the incidence of pemphigus vulgaris is within the range of the previous papers, and the sex distribution is also similar.[8-10] The increasing trend in incidence of bullous pemphigoid has not been previously shown. A possible explanation is increased ascertainment, but if this is the case the increasing trend has not levelled off.

One year mortality from bullous pemphigoid is 19%, intermediate between that described in the United States (11%) and Europe (25-41%).[2 4-6 13] One year mortality for pemphigus vulgaris is higher than that described in a study in Turkey (4.8%) but lower than that in a study from the United States (54%; mean age of cases not given).[10 11] The designs used in both of these studies are susceptible to selection bias, and drawing robust conclusions from the findings is difficult.

Interpretation and implications

Our study shows that rates of bullous pemphigoid seem to be increasing over time. Rates of pemphigus vulgaris are increasing to a lesser degree despite the facts that both disorders are more common with increasing age and that the multivariate analysis was adjusted for age. Possible explanations include ascertainment bias, misclassification, and a true increase in incidence. Our study has also shown high mortality associated with these diseases.

Contributors: See bmj.com
Funding: SML is funded by a grant from the BUPA Foundation. LS is supported by a Wellcome Trust senior research fellowship in clinical science. JW is supported by a Department of Health clinician scientist fellowship, and KMF is also funded from that fellowship.
Competing interests: None declared.
Ethical approval: Nottingham research ethics committee.
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Adam BA. Bullous diseases in Malaysia: epidemiology and natural history. *Int J Dermatol* 1992;31:42-5.
2 Garcia-Doval I, Mayo E, Nogueira Farina J, Cruces MJ. Bullous pemphigoid triggered by influenza vaccination? Ecological study in Galicia, Spain. *Br J Dermatol* 2006;155:820-3.
3 Jung M, Kippes W, Messer G, Zillikens D, Rzany B. Increased risk of bullous pemphigoid in male and very old patients: a population-based study on incidence. *J Am Acad Dermatol* 1999;41:266-8.
4 Gudi VS, White MI, Cruickshank N, Herriot R, Edwards SL, Nimmo F, et al. Annual incidence and mortality of bullous pemphigoid in the Grampian region of north-east Scotland. *Br J Dermatol* 2005;153:424-7.
5 Colbert RL, Allen DM, Eastwood D, Fairley JA. Mortality rate of bullous pemphigoid in a US medical center. *J Invest Dermatol* 2004;122:1091-5.
6 Roujeau JC, Lok C, Bastuji-Garin S, Mhalla S, Enginger V, Bernard P. High risk of death in elderly patients with extensive bullous pemphigoid. *Arch Dermatol* 1998;134:465-9.
7 Hietanen J, Salo OP. Pemphigus: an epidemiological study of patients treated in Finnish hospitals between 1969 and 1978. *Acta Derm Venereol* 1982;62:491-6.
8 Pisanti S, Sharav Y, Kaufman E, Posner LN. Pemphigus vulgaris: incidence in Jews of different ethnic groups, according to age, sex, and initial lesion. *Oral Surg Oral Med Oral Pathol* 1974;38:382-7.
9 Salmanpour R, Shahkar H, Namazi MR, Rahman-Shenas MR. Epidemiology of pemphigus in south-western Iran: a 10-year retrospective study (1991-2000). *Int J Dermatol* 2006;45:103-5.
10 Uzun S, Durdu M, Akman A, Gunasti S, Uslular C, Memisoglu HR, et al. Pemphigus in the Mediterranean region of Turkey: a study of 148 cases. *Int J Dermatol* 2006;45:523-8.
11 Ryan JG. Pemphigus: a 20-year survey of experience with 70 cases. *Arch Dermatol* 1971;104:14-20.
12 Lewis JD, Schinnar R, Bilker WB, Wang X, Strom BL. Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. *Pharmacoepidemiol Drug Saf* 2006;16:393-401.
13 Joly P, Roujeau JC, Benichou J, Picard C, Dreno B, Delaporte E, et al. A comparison of oral and topical corticosteroids in patients with bullous pemphigoid. *N Engl J Med* 2002;346:321-7.

Accepted: 7 May 2008

Drips, drains, and dressings

After a restless night at home with increasing abdominal pain, followed by bilious vomiting in casualty and an ultrasound scan that showed "a thick-walled gall-bladder with multiple stones, some impacted in the neck," the diagnosis of cholecystitis was obvious.

So, my infrequent episodes over 20 years of retrosternal pains and unrelated rigors diagnosed by me as due to oesophageal spasm and viral infections had been nothing of the sort. How unsurprising—doctors who treat themselves have fools as patients.

With my minimal right upper quadrant tenderness and a neutrophil count of 14 000, treatment was also obvious: intravenous antibiotics and remove the offending organ in six weeks' time.

An unpleasant week in hospital followed, largely because of the illness but compounded by poor communication, perfunctory examinations, and a misleadingly normal temperature chart (but then temperature goes up after a rigor and not at the time).

It was great to go home, but five days later, after hours of rigors, I was back in casualty on a Friday night. This time, a thorough history and examination was conducted by the on-call surgical registrar. My minimal abdominal tenderness and neutrophil count of 21 000 resulted in intravenous antibiotics being recommenced and, in the morning, a visit from the on-call consultant.

Like his registrar, the consultant had plenty of time to take a full history and make a thorough examination. Some fullness over the gallbladder area was noted, and the first thought was to continue intravenous treatment—but wait. Rigors a few days after antibiotics, a rising neutrophil count, and some abnormality on examination led to a rethink: "Well, you don't look ill enough and should have more tenderness to have an empyema, but how else to explain the developing picture?" And thank heavens for that time honoured surgical aphorism, "If in doubt, cut it out."

The findings: gallbladder stuck to colon and duodenum, thick walled in part, but thin and friable in others, containing multiple stones, some impacted in the neck—and containing 150 ml of pus.

So perhaps for me, postponement of a failed interview with the heavenly choir. And for all of us; don't be slavishly led by test results, but allow them to do their job in building up the complete clinical picture. Thank you to the on-call surgeons for taking the time to listen.

Anthony Craighill GP principal, Shoreham, Sevenoaks
susan.craighill@btinternet.com

Cite this as: *BMJ* 2008;337:a98