



This is an abridged version; the full version is on bmj.com

Editorial by Clauw and p 1373

Gulf War Illnesses Research Unit, Department of Psychological Medicine, Guy's, King's, and St Thomas's School of Medicine, London SE5 8AZ

Matthew Hotopf reader

Anthony S David professor

Lisa Hull research assistant

Vasilis Nikalaou statistician

Catherine Unwin study coordinator

Simon Wessely professor

Correspondence to: M Hotopf m.hotopf@iop.kcl.ac.uk

BMJ 2003;327:1370-2

Gulf war illness—better, worse, or just the same? A cohort study

M Hotopf, Anthony S David, Lisa Hull, Vasilis Nikalaou, Catherine Unwin, Simon Wessely

Abstract

Objectives Firstly, to describe changes in the health of Gulf war veterans studied in a previous occupational cohort study and to compare outcome with comparable non-deployed military personnel. Secondly, to determine whether differences in prevalence between Gulf veterans and controls at follow up can be explained by greater persistence or greater incidence of disorders.

Design Occupational cohort study in the form of a postal survey.

Participants Military personnel who served in the 1991 Persian Gulf war; personnel who served on peacekeeping duties to Bosnia; military personnel who were deployed elsewhere ("Era" controls). All participants had responded to a previous survey.

Setting United Kingdom.

Main outcome measures Self reported fatigue measured on the Chalder fatigue scale; psychological distress measured on the general health questionnaire, physical functioning and health perception on the SF-36; and a count of physical symptoms.

Results Gulf war veterans experienced a modest reduction in prevalence of fatigue (48.8% at stage 1, 43.4% at stage 2) and psychological distress (40.0% stage 1, 37.1% stage 2) but a slight worsening of physical functioning on the SF-36 (90.3 stage 1, 88.7 stage 2). Compared with the other cohorts Gulf veterans continued to experience poorer health on all outcomes, although physical functioning also declined in Bosnia veterans. Era controls showed both lower incidence of fatigue than Gulf veterans, and both comparison groups showed less persistence of fatigue compared with Gulf veterans.

Conclusions Gulf war veterans remain a group with many symptoms of ill health. The excess of illness at follow up is explained by both higher incidence and greater persistence of symptoms.

Introduction

Consensus exists that service in the 1991 Gulf war resulted in increased symptomatic ill health. We know of no studies on the prognosis among Gulf war veterans. In 1997 we compared a large random sample of members of the armed forces who served in the 1991 Gulf war,¹ with two military control cohorts. This study

assesses the outcomes of these cohorts four years later. Our two main aims were to compare the prevalence of various health outcomes over time and between cohorts, and to determine rates of incidence and remission for clinically important fatigue and psychological distress after adjusting for potential confounders.

Method

Participants

Our original study consisted of three groups: personnel who served in the Gulf war between 1 September 1990 and 30 June 1991 (the Gulf cohort); personnel who served on UN peacekeeping duties in Bosnia between 1 April 1992 and 6 February 1997 (the Bosnia cohort); and personnel who were serving in the armed forces on 1 January 1991 but who were not deployed to the Gulf (the "Era" cohort).¹ We took a random sample of all Gulf veterans, with oversampling of women. Sampling of the other two cohorts was frequency matched in terms of sex, age, reservist status, officer status, service, and a measure of fitness.

We used random stratified sampling to select respondents from stage 1 into the present study. All women were selected. We stratified the sampling on the severity of fatigue at stage 1. The selection process included all male veterans with a fatigue score greater than 8 (511 Gulf, 115 Bosnia, and 120 Era); for Gulf, a 50% sample of veterans with fatigue scores of 4-8, 484 Gulf and all 333 Bosnia and 364 Era with these scores, and an approximately one in eight sample of veterans with fatigue scores less than 4 in order to represent asymptomatic individuals (n = 250 in each group).

We used three mailings, tracing non-responders via the NHS central registry, the online electoral registry, service pension, discharge sources, and the UK Department of Social Security.

Questionnaire and outcomes

The questionnaire included a fatigue scale²; the 12 item general health questionnaire (a screening questionnaire for common mental disorders)³; the SF-36 instrument for physical health and functional capacity; and a list of 50 common symptoms. We defined cases of fatigue as having a score on the fatigue scale of greater than 3 and cases of psychological distress as having a score greater than 2 on the general health questionnaire. We defined cases

Table 1 Prevalence of categorical outcomes in the three cohorts. Values are percentages (95% confidence intervals) unless otherwise indicated

	Gulf			Bosnia			Era		
	Stage 1	Stage 2	Ratio* (new cases/recovered cases)	Stage 1	Stage 2	Ratio* (new cases/recovered cases)	Stage 1	Stage 2	Ratio* (new cases/recovered cases)
Fatigue cases	48.8 (45.4 to 52.2)	43.4 (39.9 to 46.8)	0.65 (0.45 to 0.85)	29.0 (25.6 to 32.4)	32.7 (28.6 to 36.8)	1.21 (0.83 to 1.59)	22.8 (20.0 to 25.6)	22.0 (18.6 to 25.4)	0.91 (0.56-1.26)
Post-traumatic stress reaction cases	12.4 (10.7 to 14.2)	10.8 (9.1 to 12.5)	0.73 (0.47 to 0.99)	5.7 (4.0 to 7.4)	6.0 (4.2 to 7.8)	1.07 (0.49 to 1.65)	4.0 (2.6 to 5.3)	6.6 (4.8 to 8.4)	2.45 (0.88-4.02)
General health questionnaire cases	40.0 (36.8 to 43.2)	37.1 (33.8 to 40.4)	0.79 (0.59 to 1.00)	29.2 (25.5 to 32.9)	31.5 (27.4 to 35.6)	1.25 (0.84 to 1.67)	25.3 (21.7 to 28.9)	23.8 (20.1 to 27.6)	0.88 (0.56-1.20)
Self reported Gulf war syndrome	18.6 (16.2 to 21.1)	15.8 (13.3 to 18.2)	0.58 (0.25 to 0.90)						

All prevalence estimates are weighted for sampling.

*Values of <1 indicate declining prevalence. Ratios are weighted for sampling.

Table 2 Scores (95% confidence intervals) for continuous measures by cohort and stage

	Gulf			Bosnia			Era		
	Stage 1	Stage 2	Difference	Stage 1	Stage 2	Difference	Stage 1	Stage 2	Difference
SF-36* physical function	90.3 (88.3 to 91.3)	88.7 (87.6 to 89.9)	-1.6 (-2.5 to -0.7)	95.4 (94.4 to 96.4)	92.9 (91.6 to 94.1)	-2.6 (-3.8 to -1.3)	92.1 (90.6 to 93.6)	90.8 (89.2 to 92.3)	-1.3 (-2.7 to 0.1)
SF-36* health perception	65.8 (64.1 to 67.5)	65.9 (64.2 to 67.6)	0.1 (-1.2 to 1.4)	76.2 (74.4 to 77.9)	72.9 (71.0 to 74.8)	-3.3 (-5.1 to -1.6)	76.8 (75.0 to 78.6)	74.4 (72.4 to 76.4)	-2.4 (-4.2 to -0.6)
General health questionnaire	14.5 (14.1 to 14.9)	14.2 (13.8 to 14.5)	-0.3 (0.1, -0.6)	13.1 (12.7 to 13.6)	13.2 (12.7 to 13.7)	0.1 (-0.4 to 0.6)	12.4 (12.0 to 12.8)	12.9 (12.5 to 13.3)	0.5 (0.05 to 1.0)
Fatigue	17.8 (17.4 to 18.1)	16.9 (16.5 to 17.2)	-0.9 (-1.2 to -0.6)	15.6 (15.2 to 16.0)	15.3 (14.9 to 15.7)	-0.3 (-0.7 to 0.2)	14.7 (14.3 to 15.0)	14.9 (14.5 to 15.3)	0.2 (-0.2 to 0.6)
Total symptoms	11.0 (10.4 to 11.6)	10.7 (10.1 to 11.3)	-0.3 (-0.8 to 0.1)	6.2 (5.6 to 6.8)	7.9 (7.3 to 8.5)	1.7 (1.2 to 2.3)	5.3 (4.8 to 5.8)	6.4 (5.8 to 7.0)	1.1 (0.6 to 1.6)

All scores are weighted for sampling.

*SF-36 scales range from 0-100, with higher scores indicating better health.

For SF-36 scores, negative differences in mean indicate a worsening in health. For other scales, negative scores indicate an improvement in health.

of “stress reaction” from a checklist of symptoms described in previous work.¹

Statistical analyses

Response bias—We defined four groups—responders, “refusers,” “returns to sender,” and “no information,” and compared their frequency by cohort, and also compared demographic variables and stage 1 health outcomes across cohorts, using Scheffé’s test.⁴

Follow up health outcomes—We calculated the prevalence of binary outcome variables and present these in relation to baseline scores. For binary outcomes we present the matched odds ratio, which is the proportion of incident cases to recovered cases for each outcome. For continuous outcomes we present stage 1 and 2 scores and mean differences.

Results

Response rates

The response rate for those eligible to receive a questionnaire was 71.6%. The response rate was higher in the Gulf cohort than in the other two cohorts ($P=0.03$). There were similar types of non-respondents in Gulf and Bosnia, but the Era group had a higher proportion of refusers than the other two cohorts. Response rates were lower in men, younger participants, and those who were unmarried. Non-responders rated their health as poorer at stage 1 for physical disability and general health perception but were less likely to have been cases on the general health questionnaire.

The sociodemographic characteristics of Gulf and Era were broadly similar. The Bosnia group were

younger, less likely to be married, more likely to have remained in service, and only from the Army.

Table 1 shows the prevalence of categorical outcomes at stages 1 and 2. Gulf veterans had higher rates than the other two cohorts, and this difference is maintained between stages 1 and 2. For Gulf we found a modest reduction in the prevalence of fatigue, post-traumatic stress reaction, general health questionnaire cases, and self reported “Gulf war syndrome.” For Bosnia and Era we found no changes other than an increase in the prevalence of post-traumatic stress reaction in the Era group, which was not significant ($P>0.05$).

Table 2 shows the scores for continuous measures. The Gulf cohort was less healthy than the other two cohorts at both stages. A decline in physical functioning affected each of the three cohorts (non-significant for Era). Health perception declined for both Bosnia and Era but not for Gulf. The Gulf veterans showed a modest reduction in fatigue scores and non-significant reductions in general health questionnaire scores and total symptoms. The other two cohorts showed a general tendency to experience more symptoms over time.

We explored the possibility that differences in prevalence between cohorts could have been explained by either higher incidence, or greater persistence, of symptoms. The incidence risk for fatigue and general health questionnaire caseness was lower in Era than the other two cohorts. Controlling for stage 1 sociodemographic variables reduced the differences, but the Era group remained less likely to experience new fatigue than the Gulf group. The Gulf

group were more likely to experience persistent fatigue compared with the Era and Bosnia cohorts, an effect that remained significant after controlling for potential confounders ($P = 0.009$).

Discussion

Main findings

Gulf war veterans continue to experience symptoms that are considerably worse than would be expected in an equivalent cohort of military personnel. However, Gulf war veterans are not deteriorating and do not have a higher incidence of new illnesses.

Gulf war veterans show a disappointing stability in the prevalence of the main disorders we studied. Although the prevalence of the symptom based disorders lessened for Gulf veterans, physical functioning and health perception measured on the SF-36 barely changed. The reduced physical functioning may have been due to increasing age. The two comparison groups had some worsening in health on the SF-36 scales and more physical symptoms. This implies that some worsening of these health outcomes is expected over time, presumably due to advancing age. Overall the Gulf veterans continued to experience poorer health on all measures.

We also examined whether the raised prevalence in Gulf veterans was explained by a greater incidence of disorders or more persistence. No clear pattern emerged. We found some evidence that the incidence of fatigue and caseness on the general health questionnaire was higher in the two deployed groups (Bosnia and Gulf) than in Era, but the latter was explained by confounding. Fatigue was more likely to be persistent for Gulf veterans.

Limitations of the study

The follow up rate of just over 70%, leaves room for bias. Because follow up was worse in participants with poorer health at stage 1 we have probably slightly underestimated the prevalence of the disorders under study. This bias seems to have been similar across cohorts, and is unlikely to have changed the main findings of this study. We measured health on a range of self report items, which are open to reporting biases.

Implications

The nature of Gulf war illness remains ambiguous. If the illness represented the prodrome of a known disease this has yet to declare itself.^{5,6} We think that the non-specific increase in symptoms reported by our and other studies is likely to remain poorly understood in terms of conventional biomedical diseases. As time passes it becomes increasingly difficult to find causes of illnesses in veterans of the 1990-1 Gulf war. We suspect that different psychosocial, military, and environmental risk factors may determine onset and recovery.

We thank Nick Blatchley and Simon Satchell from the Gulf Veterans Illness Unit of the Ministry of Defence for assistance in tracking participants, Michael Dewey for statistical advice, and the participants for their patience in once again completing lengthy questionnaires.

Contributors: See bmj.com

What is already known on this topic

Veterans of the 1990-1 Gulf war experience poorer health on most subjective outcomes than non-deployed military personnel

No satisfactory follow up studies have assessed outcome of veterans of the Gulf war over more than one wave of data collection, so it is unclear whether veterans are getting worse, staying the same, or getting better

What this study adds

Gulf war veterans still have considerably poorer subjective health than appropriate military controls

The health of Gulf war veterans has improved, but this improvement is relatively minor

For comparison groups there has been a worsening of health on some outcomes, which is probably due to ageing

The health gap between Gulf war veterans and comparison groups has therefore narrowed slightly

Funding: US Department of Defense, UK Medical Research Council.

Competing interests: None declared.

Ethical approval: The study received approval from the relevant research ethics committees.

- 1 Unwin C, Blatchley N, Coker W, Ferry S, Hotopf M, Hull L, et al. Health of UK servicemen who served in the Persian Gulf war. *Lancet* 1999;353:169-78.
- 2 Chalder T, Berelowitz C, Pawlikowska T. Development of a fatigue scale. *J Psychosom Res* 1993;37:147-54.
- 3 Goldberg D, Williams P. *A users' guide to the general health questionnaire*. Windsor: NFER-Nelson, 1988.
- 4 Scheffé H. A method for judging all contrasts in the analysis of variance. *Biometrika* 1953;40:87-104.
- 5 Sharief M, Pridden J, Delamont R, Unwin C, Rose M, David A, et al. Neurophysiological evaluation of neuromuscular symptoms in UK Gulf war veterans. A controlled study. *Neurology* 2002;59:1518-825.
- 6 David A, Farrin L, Hull L, Unwin C, Wessely S, Wykes T. Cognitive functioning and disturbance of mood in UK veterans of the Persian Gulf war: a comparative study. *Psychol Med* 2002;32:1357-70. (Accepted 14 October 2003)

Medscape conference coverage

American College of Gastroenterology 68th Annual Scientific Meeting and Postgraduate Course, Baltimore, Maryland

- Irritable bowel syndrome/other functional gastrointestinal disorders
- Acid related disorders/advances in endoscopy
- Inflammatory bowel disease/colorectal cancer and liver disease
- Coeliac disease
- Obesity for the gastroenterologist

Coverage of these sessions is available, free, on the *BMJ's* website: <http://bmj.com/medscape/gastroenterology/acg1/>