

Self-Reported Health Status and Presentations to a Warship's Sickbay¹

by
M.J. Stone ²

Abstract

This paper reports on a feasibility study that examined the self-perceived mental health status of individuals in a Royal Australian Navy (RAN) sea-going population. Sixty-three members of a convenience sample (the ship's company of HMAS PERTH) completed the General Health Questionnaire (GHQ) 30-item version and a "Naval Health Survey" (NHS). Using a cross-sectional survey design with longitudinal follow-up, paired data were generated. The GHQ yielded a rise from 38% to 44% over the study period. Formal psychological/psychiatric presentations to the sickbay represented less than 1% of all presentations. Despite not reaching statistical significance (at the 0.05 level), elevations of GHQ and NHS scores for the sample noted over the study period suggest a fall in mental wellbeing. A subgroup analysis showed some important differences in self-perceived health status amongst sea-going naval personnel over a period of time at sea. The direction and magnitude of this change varied between subgroups characterized by rank and rate.

The Study

The study was conducted in HMAS PERTH, a RAN DDG warship, where the author served as medical officer for five months. The rationale for focusing on mental health is that, from the outset, the Navy's sea-going population is deemed "fit for sea". While physical fitness is relatively guaranteed, the less explored and often less volunteered mental wellbeing of individuals is equally important in assessing the health of members of sea-going populations.

The study's aims were, therefore, to:

- a. examine the self-reported health of the crew, specifically their psychological wellbeing, using survey questionnaires;
- b. explore the effect of a period at sea on these measures;
- c. correlate outcomes of the GHQ-30 with a Naval Health Survey (NHS);
- d. determine if there were any high risk subgroups for psychological illness; and
- e. examine if any change in self-reported health status is reflected in the number and type of presentations to the ship's sickbay.

Fundamental to mental health in occupational settings is that psychological illness may impact on the effectiveness of an individual. In military settings, and specifically the remote naval environment, members may occupy highly specialised and integral roles. Therefore, psychological illness can have a significant and dramatic impact on the operational effectiveness of a deployed unit. Furthermore, it is widely recognised that psychological disturbance in community and occupational settings may be transient and go undetected and untreated by health care workers.¹ The obverse may equally be true. It has been suggested that stress in the workplace may be a causal factor of psychiatric illness²).

HMAS BRISBANE, a similar DOG.
Photograph: Courtesy CMDR N. Westphalen



Background

Several studies have investigated psychological distress and psychosomatic complaints in military settings.^{3.4.5.6.7.8} Military personnel are often required to work in austere conditions. Emotional stressors, such as disruption of normal sleep patterns, lack of usual home comforts, harsh environmental conditions, unappetising meals, fear of the unknown, boredom, and separation from family and friends, have been cited in a study on soldiers deployed in Operation Desert Shield.⁹ Also, Burr, Woodruff and Banta, determined that U.S. Navy personnel, during at-sea operations in the Persian Gulf, experienced psychological fatigue, confusion, tension, anxiety, depression and degradations in physical health.⁴ It is important to appreciate that these U.S. Navy personnel were in a conflict zone at the time, and therefore likely to be under more duress than the sample in this study. In the civilian sphere, similar environmental hardships at sea of excess heat, humidity and noise; broken rest and poor sleep, long working hours, separation from family and home, exposure to bad weather and the stress of meeting the pressures of an everchanging industry have also been identified.¹⁰

Health Status Surveys

A literature search of health status surveys yields a myriad of different instruments. A critical review of these surveys, particularly the Duke Health Profile and the 36-item Short- Form Health Survey, revealed a focus on negative aspects of health and a preponderance of items which were largely irrelevant to a physically fit, mobile population.^{11.12.13}

General Health Questionnaire

The GHQ, initially developed by Goldberg in 1972, was designed to be a self-administered screening tool to detect non-psychotic psychiatric disorders in non-psychiatric settings, such as community groups, and in primary care.¹⁴ As a screening instrument, the GHQ looks at the "hinterland" between psychological illness and psychological wellbeing. It is concerned with breaks in normal function and recognises that psychological disturbance may be transient and spontaneously remit without professional intervention. The GHQ does not attempt to detect lifelong psychological phenomena, such as personality disorders, and is primarily concerned with the appearance of new symptoms that are of a distressing nature to the individual.¹⁴

The GHQ in 12, 28, 30 and 60 number-of- item formats has been widely validated across a variety of cultures and settings.^{1.15.16.17} In military settings the GHQ, in various number- of-item formats, has been used both as a survey instrument and for assessing individuals in more clinical environments.^{5.6.7.18}

Despite the GHQ's use in a variety of settings, it has not been validated or undergone reliability studies for the population under study. It cannot be assumed to provide a high level of criterion validity in this study. This measure would have been impractical and unethical given the closed environment nature of the study and the medical officer being the researcher.

Goldberg and Williams in their review of a number of studies report that there is conflicting evidence for the role of gender in affecting GHQ scores.¹⁴ This is clearly a consideration in interpreting the results of this study, having been carried out on an all-male population.

Other demographic variables, of marital status and employment status, have been shown to have an effect with increased scores among unemployed persons, and divorced and separated women. Married men have been noted to report particularly low scores in general community samples.¹⁴

Method

The study used a quasi-experimental cross-sectional survey design with longitudinal follow up on a convenience sample. There was no control group.

Each survey consisted of the GHQ plus the NHS. In addition, the first survey collected demographic details of age, rank, rating, marital status. Sea-posting status, time in a defence force, and whether or not the member lived onboard the ship when it was alongside in homeport. Change in score was used as the dependent variable,

with the aforementioned variables and initial score entered as independent variables in regression analyses. Sickbay presentations were considered in descriptive analysis only.

The first survey was administered on the first day on leaving the home port after the ship had been alongside for 14 days. Two nine-day (inclusive) periods at sea and a three-day period alongside ensued before respondents to the first survey were asked to fill in the second. The first nine-day period was spent in cruising watches and the second in defence watches.

Subjects

The ship's company, excluding the commanding officer and the author totaled 328 at the commencement of the study. The median age was 25.0 years. The study was approved by the Australian Defence Medical Ethics Committee (ADMEC) and Curtin University Postgraduate Ethics Committee. Participation in the study was entirely voluntary. A covering letter and copy of "Guidelines for Volunteers" from ADMEC explained participants rights as a volunteer and stressed that the member was free to withdraw from the study at any time with no prejudice or effect on future medical care or career.

All members except the commanding officer were invited to participate in the study. Contact was made at main messing times for both phases of the survey. Ninety-five (29%) volunteered to participate and returned completed questionnaires during the first phase of the survey and 63 (19.2%) were entered for analysis at the completion of data gathering.

NHS

Due to the uniqueness of the study population, a questionnaire with more questions relevant to living in a naval military environment was thought to be important. Questions were presented in similar format to improve the flow of the questionnaires. It was hoped the NHS would be perceived to be relevant to the target population and therefore promote involvement in the study.

Informal discussion by the author with sailors indicated that many were concerned with the possible health effects of specific environmental hardships, such as excessive noise and heat, exposure to smoke and fumes (especially related to explosives and gun firing), and radiation exposure. The NHS was constructed not to inquire specifically of these stresses but to explore the possibility that they may contribute to disturbances in sleep, energy, concentration, self-confidence and overall sense of wellbeing.

The completed survey format, including the two questionnaires and demographic items, was presented to a group of uniformed personnel in a pilot setting and was subject to peer review by other military medical officers for face and content validity.

Scoring

The GHQ has been reported in the literature as being able to be scored in three possible ways. The first is the GHQ method in which each item is scored "0 0 1 1".¹⁴ This method eliminates "middle" scoring. Another method uses a Likert scale which assigns scores as "0 1 2 3" for each of the items. This has been shown to yield less skewed curves.¹⁴ A final means is the *cGHQ* method, originally described by Goodchild and Duncan-Jones.¹⁵ This method gives a score for those replying "same as usual" for a negative item as this implies some degree of chronicity.

Such scoring has been shown to yield overall scores that follow a more normal distribution and, in Goodchild and Duncan-Jones' study, showed improved sensitivity for psychiatric caseness as determined by the Present State Examination.^{14,19} Desirable threshold or cut scores for psychiatric "caseness" for the GHQ scored in GHQ mode have been reported between 2/3 to 12/13.¹⁴ A cut score of 4/5 is most often recommended but this is dependent on the GHQ having been validated in the population of interest.

Early studies involving the GHQ-30 in Australia, England and the USA indicate a cut score of 4/5 to yield optimum sensitivity and specificity for caseness across a wide variety of populations.^{1,15,20,21,22} The question, "Been getting out of the house as much as usual?" was not relevant to this study and was removed from further

analysis. Thus, all scores are out of 29 and not 30 questions, and a cut score of 5 was used in determining "high" scorers on the first phase of the survey.

Sickbay presentations

All presentations to the sickbay were recorded. Patients presented to the sickbay onboard PERTH for 29 days out of the total 35 days over which the study was conducted. This period includes 1 day before deployment. The reason for the omitted days of recording was two weekends (including a three-day) before deployment and one day, when alongside in Guam, where another ship provided medical guard.

Patients were listed according to their provisional diagnosis as given by medical officer or medic. These diagnoses were later reviewed and classified into one of 14 diagnostic categories. Representations of the same problem, routine follow-ups and reviews were not recorded. New presentations only were listed for analysis. A patient could present with one or more separate problems and could be listed under one or more diagnostic categories. The diagnostic category list was based on a historical review of presentations to the sickbay on the ship over the preceding few months. Most categories were anatomically based and designed to minimise inter-rater variability and misclassification bias.

Statistical analysis

Paired samples t-tests (PSTT) were used to compare the scores (both for GHQ and NHS, scored in Likert, GHQ and cGHQ) of respondents over the two different time periods. Since the study utilised a convenience sample the assumption of random sampling of differences was assumed but could not be ensured. Assumptions of independent observations (for one of the variables), normality (for one of the variables) and constant variance were met.

A stepwise method was used in the regression analysis. Chi-square tests of hypothesised proportions are made on some of the category presentations about day 19. Chi-square test stipulations of non-biased sampling, independent observations, mutually exclusive categories that include all of the observations, and sufficiently large expected frequencies, are assumed.²³

Results

Initial and change in self-reported psychological wellbeing.

In interpreting the scores, no assumptions were made on the interval magnitude of the score reflecting severity of psychological distress. If the cut score of five for GHQ(GHQ) scores is accepted, then it does provide an indicator of prevalence of psychological distress. In other words, if the GHQ scored in GHQ mode was being used as a screening instrument, then those scoring five or higher would be flagged as "cases" and be referred for psychiatric assessment.

Thirty-nine (41.1%) of the original 95 participants scored GHQ(GHQ) scores of five or greater. This is a disturbingly high proportion.

Of the sample of 63 persons, twenty-four or 42.9% scored five or higher on initial GHQ(GHQ) score. On follow up survey, 28 out of the 63 (44.4%) scored five or higher. A chi-square test of hypothesized proportions yields a chi-square value between 0.5 and 0.6 and thus the null hypothesis of the same population proportion scoring GHQ(GHQ) five or higher over the two surveys, cannot be rejected.

The NHS correlated (linear) closely with the GHQ when scored in Likert mode (Pearson correlation = 0.8). This is not interpreted to mean that the two questionnaires were measuring the same thing but suggests that the NHS scored in Likert is measuring some psychological distress parameter. Scored in Likert, the NHS yielded an approximately normal curve for distribution of scores, a desirable attribute for a screening questionnaire. Other score distributions for the NHS do not correlate as closely as the Likert scores and no attempt was made to use NHS scores to determine a cut score for indicator of psychological illness. This again would have necessitated the use of a validating psychiatric interview.

GHQ responses scored in all three modes of Likert, GHQ and cGHQ, returned difference distributions that, while not meeting formal Kolmogorov-Smirnov tests of normality, were not so non-normal that the central limit theorem could not be applied. Thus, they were considered approximately normal for the purposes of calculating paired sample confidence intervals for differences between means. NHS difference scores distributions were also considered normal in further analyses. Normality of score distributions for original paired samples is not important as no assumptions are made about the ship's crew as a population from which the final paired sample is taken.

Overall, the study found mean differences between the paired scores for the final sample and individual subgroups were generally small but some interesting differences were observed.

In examining the regression analyses, performed using a stepwise regression method, the most notable outcome is that for all difference scores (scored in Likert, GHQ and cGHQ). Entered as the dependent variable, only the first phase score reached significance to be entered into the regression equation.

Due to the relatively small sample size the categorical variables rank, rate and marital status were not converted to binary variables and instead were considered in separate subgroup paired sample T-test analyses.

In considering the results of PSTTs for the final sample of 63 responders, mean differences for the six pairs of scores show that scores were higher on the second phase of the survey resulting in a negative value for the mean difference. However, only the NHS Likert scored reached statistical significance at the 0.05 level (95% CI -3.80, -0.01; $P = 0.049$). The mean difference measured - 1.90. The GHQ scored in Likert returned a mean difference of -2.78 ($P = 0.093$). Validation studies that have reported optimum threshold or cut scores have used the GHQ scoring method. When scored in GHQ mode, the GHQ returned a mean difference of just under one (- 0.937) and two-tailed statistical significance measured 0.241 (95% CI -2.51, 0.64). Statistical significance suffers due to the small size of the sample but a one-point shift to a higher (and therefore worse mental state) may be clinically significant.

When PSTT output for subgroups of high scorers (GHQ(GHQ) ≥ 5) including officers, senior sailors, junior sailors, and marine technical, electrical technical and combat systems operator sailors is examined, some distinct differences emerge. Firstly, the high scorers on initial survey showed an improvement in scores with a mean difference of 2.3 (95% CI 0.35, 4.24; $P = 0.022$). Scores on second survey were lower, indicating less psychological distress at that time. Whilst aggregate measures showed a decrease in mean second survey scores for this subgroup, it is important to point out that there were more high scorers (28 versus 24) at the time of the second survey.

The other subgroup that showed lower GHQ(GHQ) scores (increased sense of wellbeing) was marine technical sailors with a mean difference of 4.10 (95% CI 1.15, 7.05; $P = 0.012$).

Sickbay presentations

An examination of the type and number of sickbay presentations contributes, in descriptive terms, some understanding of the way members of the ship's company externalised their sense of wellbeing and presented to the sickbay with a problem.

The total number of presentations over the study period is depicted in Figure 1. Figure 2 shows the breakdown of presentations. There is the suggestion of increased numbers of new presentations, particularly in the final 10 days. This corresponds to the period when the ship practiced defence watches. This time represented overall longer working hours for all of the crew. Importantly, this also included the medical department and meant that the sickbay was manned constantly and members were welcome to attend at any time. The nature of the presentations at this time were largely made up of ENT, respiratory tract and general infectious presentations with a lesser but still pronounced increase in musculoskeletal and traumatic injuries.

The number of psychiatric presentations was low. One member was diagnosed with a mild depressive illness and another manifested a schizophreniform psychosis and was one of the two aeromedical evacuations conducted over the study period.

There was a relative higher number of dermatological presentations on days 20 and 21. Conditions included non-specific dermatitis (considered to be heat-related and seen most commonly in MT sailors), tinea pedis, tinea corporis, tinea versicolor and foot blisters. Conditions at this time were hot and humid. An inspection of mess decks revealed large amounts of dirty clothes and towels that were well worn, suggesting that members may have been running low on clean socks, underpants and towels. In addition, a common practice was to pin shower footwear (plastic sandals, thongs or equivalent) to towels on lockers to dry and to make secure for sea. These factors may have contributed to the increase in presentation of dermatological problems, predominately cutaneous fungal infections.

It is interesting that despite the ship remaining around Guam, and environmental conditions remaining hot and humid, dermatological complaints did not continue to rise but instead levelled off to around one to the presentations per day. This finding may indicate that changes in climatic effects may impact relatively quickly on vulnerable persons and then stabilize. It may also be possible that because tinea is common, and hence readily recognizable amongst members, many may have simply treated the condition themselves by sharing creams and powders.

ENT, general infectious and respiratory tract complaints followed a similar pattern of increased numbers of presentations from days 25 to 35. This corresponds to the second period at sea following the brief stopover in Guam. A chi-square test of hypothesized proportions about day 19 for respiratory tract complaint presentations record observed frequencies of 18 and 50, with an expected frequency of 34. This yields a chi-square value of 15.06 and statistical significance < 0.0005 ($df = 1$).

It is likely that one or more non-specific upper respiratory tract viral pathogens were contracted in Guam with person to person spread resulting in the maximum number of new presentations on days 30 to 33. A possible reason for the high numbers of general infection presentations (fevers, general feelings of being physically unwell with minimal localising signs) is that members may present very early, at the first onset of any symptoms, due to the close proximity of the sickbay and the fact that it was manned and open at all times during this period.

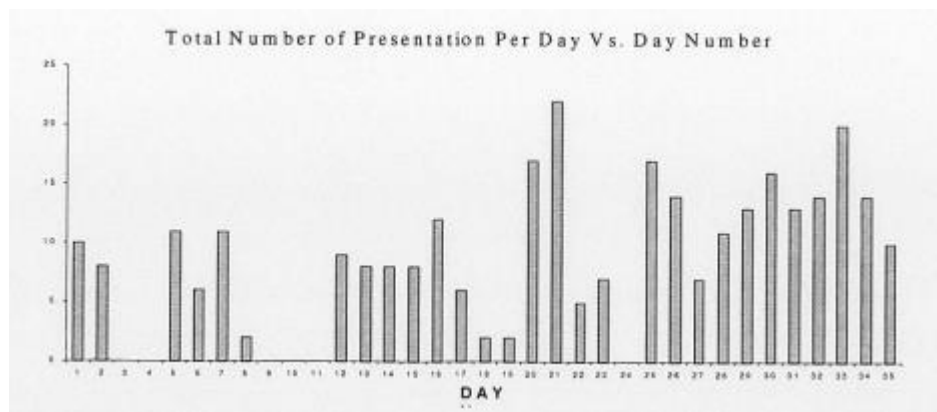


Figure 1. Total number of new presentations to the sickbay over 35 days. Days 13 to 23 and days 25 to 35 are days when the ship was at sea.

Gastrointestinal complaints were relatively infrequent and represented only one to two per day. Diagnoses included gastro-oesophageal reflux and isolated nausea, vomiting and/or diarrhoea that did not seem to be primarily motion sickness related. There were no acute infectious gastroenteritis-like illness outbreaks.

Genitourinary complaints were infrequent and totalled seven for the study period. Only one presentation was for a new sexually transmitted disease in a member who developed herpes simplex ulcers under the foreskin and secondarily developed a complication of dorsal vein thrombosis of the penis. The remainder was either known conditions such as genital warts, recurrent epididymo-orchitis or new scrotal lumps for investigation.

Motion sickness presentations were maximal on the second day of being at sea. Many of the neurological presentations (which were predominately non-specific headaches) may also represent a component of motion sickness.

Musculoskeletal, traumatic, and soft tissue infection presentations peaked on days 25 and 26, which represents the period on sailing from Guam after a short period alongside.

Most complaints were related to either sporting injuries or injuries sustained from drinking alcohol and engaging in fights or other drunken trauma. Traumatic presentations also included four bumps. All bumps occurred on day 16 and involved three MT sailors and a cook, each in separate incidents. However, even allowing for six presentations specifically related to sporting trauma after day 19 and the period in Guam, giving observed frequencies of 8 and 20 (26- 20), a statistically significant higher proportion of trauma in the second period is determined. A chi-square test of hypothesized proportions yields a chi-square value of 5.29 ($0.025 < p < 0.01$; $df = 1$).

Ophthalmological presentations were infrequent. Diagnoses included simple foreign body, uncomplicated chemical splash corneal burns and non-specific conjunctivitis.

The rest of this paper will focus on the ADF's initial efforts in seeking to address the problems of PT and sports injuries.

Table 4 provides a profile of the nature, location, and mechanism of injury for each of the leading causes of WDL and indicates the proportion of casualties associated with each.

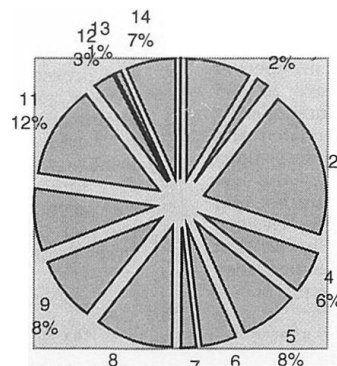


Figure 2. Presentations to sickbay over 35 days along with percentages of total number of presentations, according to diagnostic category.

- | | | | |
|----------------------|----------------------|----|-------------------|
| 1 = ENT | 5 = Musculoskeletal | 10 | General infection |
| 2 = Ophthalmological | 6 = Gastrointestinal | 11 | = Dermatological |

3= Respiratory tract	7 = Genitourinary	12	= Motion sickness
4= Neurological	8 = Trauma	13	Psychological/Psychiatric
	9 = Soft tissue infections	14	= Other

Discussion

The response rates of 29% and 19.2%, for the first and second phases of the survey respectively, were disappointingly low but better than the response rate reported by Burr, Woodruff and Banta in their cross-sectional survey study conducted on a US miscellaneous command ship of 440 persons (12.7%).⁴ Other military studies have reported surprisingly high volunteer rates such as Dahl and Kristensen's study that reported a return of 83% when they used the GHQ 12-item version to measure the psychological wellbeing of 220 Danish Army personnel.

Notable poor responders (based on rank and rate category) were seamen, marine technical sailors and boatswain's mates. Many of the other categories contained relatively small numbers and valid comparison of response rates is difficult. It is likely that, despite the study guaranteeing anonymity, potential volunteers may have been apprehensive about participating due to perceived confidentiality issues and implications of the results. In examining the other characteristics of the sample that were not able to be directly compared to the whole ship's company, such as marital status, posting status, living status and time in a defence force, the final sample closely resembles the makeup of the 95 first phase responders. Overall, the sample of 63 paired responders was considered to be representative of the ship's crew.

Despite the sample (n = 63) being a relatively young workforce population (median age = 28.0 years), the median time spent in a defence force was comparatively high (8.0 years). The mean age of 26.3 was 2.5 years younger than the mean age of the Navy population of 28.8 years (ADF Census, 1999). Responses according to rank and rate were representative of the convenient sample (N = 328) with the notable exception of seamen, marine technical and boatswains mate sailor categories, who were poor responders.

Approximately 44% of the final sample was single and this compares very closely with the total Australian Defence Force (ADF) population (44.6%). Around 40% were married and 11% were living in a de facto relationship. These measures compare closely with the total ADF population of 35.3% and 9.4% respectively, as estimated in the 1999 Australian Defence Census.²⁴ Separated and divorced personnel represented 4.8% (total ADF = 6.1%).²⁴ Twenty persons (31.7%) indicated that it was their first sea-posting, which is notable given that median time spent in a defence force was 8.0 years and mean time 9.7 years. Around one-quarter of respondents indicated they lived on board when the ship was in homeport (Sydney).

Identified high-risk subgroups

Officers and combat systems operator sailors (CSOs) recorded increased scores on the second phase of the survey indicating a decrease in their sense of wellbeing. For both of these subgroups, the mean score difference met statistical significance in all modes of scoring - Likert, GHQ and cGHQ. The mean score differences (GHQ(GHQ)) for officers and CSO sailors respectively were -2.33 (95% CI - 4.22, -0.45; P = 0.021) and -3.50 (95% CI - 5.96, -1.04; p = 0.011).

The possible reasons for these findings deserve thought. It is probable that the reasons for the difference in self-perceived health status change across MT, CSO and officer rank/rates are multifactorial. MT sailors recorded markedly lower mean GHQ(GHQ) scores suggesting a marked improvement in their mental health. It may be that MT sailors are happiest at sea performing work they are trained for - running and maintaining engines and

machinery and that this work at sea is more satisfying than work alongside. MT sailors also routinely run a watchkeeping system that is largely independent of the operational level of the ship. Combat system operator sailors and officers, on the other hand, would have varying workloads depending on the ship's operational habits. The period at sea over which this study was conducted corresponded to a busy joint exercise in which CSOs and officers' work and stress levels were likely to be high.

For the other subgroups of senior sailors, junior sailors, electrical technical sailors, those who live on board when the ship is alongside in homeport and first sea-posters, no significant changes in the score were observed. Junior sailors recorded an isolated negative score difference for NHSL of -3.03 (95% CI -5.80, -0.26; $p = 0.033$).

Married and single member marital status groups recorded universal higher scores on the second phase indicating a decrement in mental wellbeing. The differences in score were slightly higher for single members; however, neither group reached significance at the 0.05 level for any of the difference scores. Despite these findings not reaching statistical significance, they are in agreement with Goldberg and Williams report of findings from a 1987 study using the GHQ-30 of a random sample of 6498 respondents in the British Isles, in which married men were noted to have lower morbidity than single men.¹⁴

Biases

The most important potential bias in this study is the loss to follow up or non-response by 32 (33.7%) of the 95 participants, who completed a first survey but not the second for pairing. Absolute non-responders (71% of the whole ship's company) also represent bias and are unlikely to represent a random group. Non-response bias represents a potential bias for at least two reasons. The first is simply that it reduces the size of the sample and thus reduces precision for outcome estimates. The second reason lies in the potential different characteristics between responders and non-responders that may introduce bias.

It is possible that non-responders, in general, may have been experiencing higher workload levels and more psychological distress than responders. Alternatively, non-responders may have manifested a healthy psychological profile and considered the survey irrelevant and not worthy of responding to.

Every attempt was made to make contact with all of the crew but it is possible that some did not attend the 'line' at messing times by either working through two meal sittings or were otherwise indisposed when the questionnaires were distributed. While the possibility exists that subjects may have been systematically missed at survey distribution times, the author considers that this form of selection bias was minimal.

Given that the questionnaires sought response from participants with respect to the "past one to two weeks" and "recently", the potential for recall bias existed. Individuals and subgroups may have related previous (adverse) experiences from being at sea and reported these incorrectly for the period over which the study was conducted.

Every attempt was made to minimize response and reporting bias, but there is always the possibility that certain respondents may have provided false and deliberately misleading information or misinterpreted the questions. The potential for providing misleading results may also be due to evaluation anxiety and reactive effects. With respect to the reporting, recording and subsequent coding of the sickbay presentations there is the potential for observer bias. This may not only be due to perceptual differences between sickbay presenters (e.g. traumatic skin lesion versus soft tissue infection), but also inter-observer variation between the researcher and medic staff.

Confounding

Because there is no unexposed (to sea) group to compare to the exposed group, risk measures are not calculated. Further study in this area using a cohort study design with unexposed group would need to carefully randomize, stratify or match subjects to control for unmeasured variables inherent in potential confounders related to rank rate marital status, age and time in a defence force:

Limitations

The main limitation of the study relates to the lack of generalisability to other RAN sea-borne populations. The reasons for this include the uniqueness of an all-male crew and the physicality of an old steam-powered ship, the nature of which cannot be found on any other class of ship in the RAN.

Other limitations concern the internal validity of the study. With respect to the NHS, the degree of correlation between the NHS and GHQ difference scores should in no way be interpreted to mean that the two questionnaires are measuring the same phenomena. In order to validate the questionnaires an independent (preferably blinded) standardized psychiatric assessment would need to be undertaken for the study population. This was outside the cope of this study and the ethical limitation of the researcher potentially identifying participants has been discussed.

Conclusion

Overall, the final paired sample of 63 respondents was considered representative of the ship's company and this feasibility study demonstrated acceptable internal validity. The GHQ-30 needs to be validated for the RAN sea going population if it is to be further used as either a screening tool or instrument for estimating levels of psychological distress. However, the study has limited external validity and results should be cautiously interpreted in considering current application of the findings to other sea-home populations in the RAN.

The NHS did not confer any measurement benefit apart from yielding a near-normal curve when scored in Likert. This survey, again, would need to be validated if it were to be used in other studies and its further use without validation would not be helpful.

If the GHQ-30 is accepted as a valid instrument and cited cut scores for probable psychiatric illness are accepted, then the prevalence of psychiatric caseness as determined by a cut score of five for the GHQ scored in GHQ mode is disturbing. In the sample of 63 persons, prevalence of caseness rose from approximately 38% to 44% over the study period.

Changes in self-perceived health status were measured over the study period for particular subgroups but no significant change was determined for the sample as a whole. Officers and combat system operator sailors recorded a statistically significant decrease in their sense of wellbeing as determined by scores on the GHQ. In contrast, marine technical sailors recorded a statistically significant increase in their sense of wellbeing after the same period.

Sickbay presentations over the study period provided a useful descriptive measure of the type and number of health complaints. There was a low incidence of overt psychological/psychiatric presentations. Evidence was presented to support a statistically significant rise in trauma presentations for the second half of the study period when the ship observed defence watches, even when sporting and "drunken" trauma was taken into account.

The study represents a significant inroad to exploring self-reported health status in members of the RAN posted to sea-going units. Other studies need to be done on other ship populations in order to validate the findings of this study and help further understand the impact of living at sea in the naval military environment on the mental health of sea-going persons.

References

1. Tennant C. The General Health Questionnaire: A valid index of psychological impairment in Australian populations. *MJA* 1977; 2: 392-394.
2. Pflanz S. Psychiatric illness and the workplace: Perspectives for occupational medicine in the Military. *Mil Med* 1999; 164: 401-406.

3. Askenasy J, Lewin I. The impact of missile warfare on self-reported sleep quality. Part I. *Sleep* 1996; 19(1): 47-51.
4. Burr R, Woodruff S, Banta G. Associations between mood and specific health composites during U.S. Navy Persian Gulf operations. *J Psychosomatic Res* 1993; 37(3): 291-297.
5. Dahl S, Kristensen S. Health profile of Danish army personnel. *Mil Med* 1997; 162: 435-440.
6. Goyne A. Use of the GHQ in an *Australian Army population* 1998. DSPPR Research Report 4/98.
7. Johnsen B, Laberg J, Eid J. Coping strategies and mental health problems in a military unit. *Mil Med* 1998; 163(9): 599-602.
8. Johnsen L, Spinweber C. Good and poor sleepers differ in Navy performance. *Mil Med* 1983; 148: 727- 731.
9. Wasserman G, Martin B, Hyams K, Merrill B, Oaks H, McAdoo H. A survey of outpatient visits in a United States army forward unit during Operation Desert Shield. *Mil Med* 1997; 162(6): 374-379.
10. Parker A, Hubinger L, GreenS, Sargent L, Boyd R. *A survey of the health, stress and fatigue of Australian seafarers*. Canberra: Australian Maritime Safety Authority; 1997.
11. Ware J. 1995 - The status of health assessment. *Annual Rev Pub Health* 1994; 16: 327-354.
12. Parkerson G, Broadhead W, Tse C. The Duke Health Profile: A 17-item measure of health and dysfunction. *Med Care* 1990; 28(11): 1056-1072.
13. Medical Outcomes Trust. *IQOLA SF-36 Acute Australia/ New Zealand Version 1.0 -7194*. Boston: New England Medical Centre; 1994.
14. Goldberg D, Williams P. *A user's guide to the General Health Questionnaire*. Windsor: NFER- NELSON; 1988.
15. Banks M. Validation of the General Health Questionnaire in a young community sample. *Psychol Med* 1983; 13: 349-353.
16. Munoz P, Vazquez J, Pastrana E, Rodriguez F, Oneca C. Study of the validity of Goldberg's 60-item G.H.Q. in its Spanish version. *Soc Psychiatry* 1978; 13: 99-104.
17. Harding T. Validating a method of psychiatric case identification in Jamaica. *Bull WHO* 1976; 54: 225-231.
18. Epstein R, Fullerton C, Ursano R. Factor analysis of the General Health Questionnaire. *Psychol Reports* 1994; 75: 979-983.
19. Goodchild M, Duncan-Jones P. Chronicity and the General Health Questionnaire. *Bri. tJ Psychiatry* 1985; 146: 55-61.
20. Winefield H, Goldney R, Winefield A, Tiggerman M. The General Health Questionnaire: Reliability and validity for Australian youth. *ANZJ Psychiatry* 1989; 23: 53-58.
21. Jenkins R. Minor psychiatric morbidity in employed men and women and its contribution to sickness absence. *Psychol Med* 1980; 10:751-757.
22. Berwick D, Budman S, Damico-White J, Feldstein M, Klerman G. Assessment of psychological morbidity in primary care: Explorations with the General Health Questionnaire. *J Chron Dis* 1987; 40 (Suppl. 1): 71S-79S.
23. Shott S. *Statistics for Health Professionals*. Philadelphia: W.B. Saunders; 1990.
24. ADF Census. *ADF 1999 Census Statistical Tables*. Melbourne: Roy Morgan Research; 1999.
25. Watson E, Evans S. An example of cross-cultural measurement of psychological symptoms in post-partum mothers. *Soc Sci Med* 1986; 23(9): 869-874.