PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Police officers: a high risk group for the development of mental health disturbances? A cohort study
AUTHORS	Velden, Peter G; Rademaker, Arthur; Vermetten, Eric; Portengen, Marie-anne; Yzermans, Joris; Grievink, Linda

VERSION 1 - REVIEW

REVIEWER	Torbjørn Moum Professor emeritus Dept. of Behavioural Sciences in Medicine Medical Faculty, University of Oslo Norway
REVIEW RETURNED	29-Jul-2012

THE STUDY	The occupational groups selected for comparison (from a variety of studies) are extremely heterogeneous and skewed in the same direction, in particular with respect to their gender and educational composition, also to the point of not making controls for such confounders meaningful (i.e. some sort of multicollinearity may be at work). These problems become particularly accute when extremely skewed dendent variables are employed, e.g. in Table 2 in which a raw prevalence rate of .5% among psychiatric hospital employees (1 person out of a sample of 220 above the cut-off) translates into an OR which is significantly higher than that of the police officers (who exhibit higher raw prevalence rates). One would be well-adviced to drop analyses with "severe" cases as the dependent variable. Rather than using a multinomial logistic regression analyses, which in practice means postiiting occupational groups as the dependent variable, the authors should consider using the outcome variables (i.e. the SCL-90 sub-scores dichotomies) as the dependent variables (which in fact they are). Effect parameters for the putative confounders (gender and education) should be reported, and a standard binary logistic regression analysis also would allow authors to test for interaction effects between independents. The emphasis on disproving conventional wisdom regarding the stressfulness of being a police officer is somewhat displaced and not entirely convincing, but the discussion of selection effects and training are in place. The written English contains a very large number of errors, and the manuscript should be properly edited.

REVIEWER	Shu-Ling Huang Associate Professor Department of Psychology Chung-Shan Medical University Taiwan

REVIEW RETURNED	10-Oct-2012

GENERAL COMMENTS	This study explored the relative mental health of police officers compared with other occupational groups in Netherlands and to investigate if there are empirical based indications that policing can be considered as a high risk profession. The result found that officers were not more at risk for serious mental health problems than any of the other examined occupational groups in the study. However, the data of the studies adopted in this multi-comparative cross-sectional study were collected at quite different periods, a long duration of 16 years (1991-2007). Did the authors consider the changes of political, social and economic status during the 16 years? Moreover, in terms of sampling bias, were all the study groups recruited in the study representative of those occupations? In addition, one of the study groups was administered a brief scale of the SCL-90-R (17 items out of 90 items) which was not the same as the measure for the other groups. Is it possible for authors to trim data to make them comparable and reanalyze them? It would be optimal if some indicators, such as socioeconomic indexes, could be adopted to adjust potential bias from different sampling time. [Other comments] Abstract: Page 3, line 41: RET should be full-named when the abbreviation is first showed in the paper. Samples and procedures: Page 6, line 31-46: The time when Group1 and Group2 studies were conducted have not indicated. In addition, the code for study groups should be consistent in the paper. Group 1/Group 2, Police 2002/Police 2004 (Page 10, line 22-31), and Police A/Police B (in Table 2 and Table 3) all mixed up in the manuscript. Samples and procedures: Page 7, line 10-17: The numbers are not equal (1116+149≠1294). Samples and procedures: Page 7, line 41: The numbers for group 7 and group 8 in methods (281 and 238 respectively) are not consistent with the numbers in abstract for these groups (284 and 239) (Page 3, line 43). Measures: Page 8, line 34: For which year were the Dutch SCL-90-R cut-off scores for a normal population adopted in the study? The cut-off scores should be an issue particularly the data of different studies were collected at different periods. Measures: Page 8, line 46: How did authors detect the clinical and sub-clinical levels for a brief scale of the SCL-90-R? Results:

	Page 10, line 22-31: Why did the authors make ANOVA for Group1, Group2 and employees of supermarkets? Employees of supermarkets were not Group 9 who administered the brief scale of the SCL-90. Tables 2 and 3: Why were the data of group 11 not shown in tables? Instead, the results of Group 9 appeared in Tables. Did any mistake occur?

VERSION 1 – AUTHOR RESPONSE

Reviewer: Torbjørn Moum
Professor emeritus
Dept. of Behavioural Sciences in Medicine Medical Faculty, University of Oslo Norway

The occupational groups selected for comparison (from a variety of studies) are extremely heterogeneous and skewed in the same direction, in particular with respect to their gender and educational composition, also to the point of not making controls for such confounders meaningful (i.e. some sort of multicollinearity may be at work). These problems become particularly accute when extremely skewed dendent variables are employed, e.g. in Table 2 in which a raw prevalence rate of .5% among psychiatric hospital employees (1 person out of a sample of 220 above the cut-off) translates into an OR which is significantly higher than that of the police officers (who exhibit higher raw prevalence rates). One would be well-adviced to drop analyses with "severe" cases as the dependent variable.
Rather than using a multinomial logistic regression analyses, which in practice means postiiting occupational groups as the dependent variable, the authors should consider using the outcome variables (i.e. the SCL-90 sub-scores dichotomies) as the dependent variables (which in fact they are). Effect parameters for the putative confounders (gender and education) should be reported, and a standard binary logistic regression analysis also would allow authors to test for interaction effects between independents.
The emphasis on disproving conventional wisdom regarding the stressfulness of being a police officer is somewhat displaced and not entirely convincing, but the discussion of selection effects and training are in place.
The written English contains a very large number of errors, and the manuscript should be properly edited.

Response
We have re-analyzed our data in line with this helpful comment, using standard logistic regression with symptoms as dependent variables, and group membership, age, gender and education as predictors instead of multinomial log. regression. Adjusted OR of all study variables are shown in the new tables. In all analyses both groups of officers were the reference group. As a consequence we have 2 new tables (because we had to split the lengthy tables the number is actually 2 plus 2 continued, taking 4 pages). However, the findings are not affected.

We have also provided the bi-variate Odd ratios, as requested by Prof Moum. Because we present already 4 pages of tables (table 2 and 3) and 1 page consisting of one table describing the demographics of the study groups (table 1), we have only shown the bi-variate OR and p-value in an appendix table (otherwise we would have 4 pages with tables extra). We have added this table in the appendix of the revised manuscript and will ask the editor how to handle this table (include or available on request as we suggest).

We did not examine interaction effects between group membership and demographics issues because group membership is not an interval scale, and we were not interested in interaction effects

within/between demographics.

The reviewer suggest to drop the analyses on very severe symptoms (severe cases). We must admit that we disagree on this issue. The very small numbers are not small number in themselves, but a small number within each study group. Logistic regression is suited for these kind of data. We furthermore believe that including clinically relevant symptom levels (see also below point 6 of reviewer 2) add important findings, showing differences and similarities at two symptom levels (cut-offs). We expect that if readers would read our manuscript without these analyses, it would immediately raise questions about possible differences between police and other study groups when higher (clinical levels) are taken into account and analyzed. In fact, it was our research question.

Prof. Moum wrote “The emphasis on disproving conventional wisdom regarding the stressfulness of being a police officer is somewhat displaced”. We were surprised a bit by this comment because it was not our intention to suggest that officers are hardly confronted with stressful events (i.e. potential sources of stress). In fact, in the introduction we start with depicting important sources of stress officers are often confronted with (“In other words, although police officers are more frequently confronted with critical incidents than for example employees…..”), but perhaps we misinterpret this comment.

Nevertheless, we were glad that the reviewer concluded that “the discussion of selection effects and training are in place”. This discussion is also based on the outcomes using the very severe cases.

We have furthermore improved the text in general, with the help of a native speaker.

Reviewer: Shu-Ling Huang
Associate Professor
Department of Psychology
Chung-Shan Medical University
Taiwan

General impression

This study explored the relative mental health of police officers compared with other occupational groups in Netherlands and to investigate if there are empirical based indications that policing can be considered as a high risk profession. The result found that officers were not more at risk for serious mental health problems than any of the other examined occupational groups in the study. However, the data of the studies adopted in this multi-comparative cross-sectional study were collected at quite different periods, a long duration of 16 years (1991-2007). Did the authors consider the changes of political, social and economic status during the 16 years? Moreover, in terms of sampling bias, were all the study groups recruited in the study representative of those occupations? In addition, one of the study groups was administered a brief scale of the SCL-90-R (17 items out of 90 items) which was not the same as the measure for the other groups. Is it possible for authors to trim data to make them comparable and reanalyze them? It would be optimal if some indicators, such as socioeconomic indexes, could be adopted to adjust potential bias from different sampling time.

Response
The data were obtained during the years 1991-2007. In these years society has indeed changed in several aspects, but not or hardly in other aspects.

The question that remains however, i.e. how we interpret the comment of Prof. Huang, whether societal changes may have affected mental health, and more specifically whether they affected the mental health of some specific occupational groups differently than other occupational groups: Can't we directly compare outcomes among study groups because prevalence of mental health problems were not stable across years between groups?
This is a very interesting problems, but we are not aware of any study among specific occupational groups that enables firm conclusions on this issue.
Nevertheless, epidemiological longitudinal studies among the general population on mental disorders in the Netherlands, but also in the US may help to answer this question: These studies did not find systematic increases in prevalences among the general population. With respect to the Netherlands, the NEMESIS study (cf. Vollebergh et al., 2001) showed no increases in the 12-month prevalence of any mental disorder in the general population in the period 1996 to 2007-2009, i.e. remained stable at about 17%. With respect to the US for example, the 12 month prevalence of any mental disorders was higher but also more or less stable, (29.5% in the NCS study (1990-1992 and 26.2% NCS-R study (2001-2002, Kessler et al., 2005). Of course, these studies assessed the general population, and not specific occupational groups. However, these important findings severely question whether the prevalence rates in our occupational groups really changed over time.

Nevertheless, Dr. Huang's comment made us realize that we have not properly addressed this issue in the discussion section of our manuscript (although we mentioned it briefly). For this reason we the following text in the discussion (page 14):

…… Finally, one could hypothesize that the data of our study groups were not all obtained very recently and that for instance the current prevalence of mental health problems among police officers is (much) higher. The data were in fact obtained in the period 1991-2007. It is possible in principle that the stability of the prevalence of assessed clinical and sub clinical symptoms varied across the study groups during this period. Therefor we cannot rule out the possibility that for instance mental health problems among police officers increased over these years, while those of bank employees decreased and those of soldiers remained stable. Unfortunately, we are not aware of any study assessing and demonstrating differences in trajectories of prevalence between the study groups over a period of 16 years. However, epidemiological studies among the general population examining 12-months prevalence of mental disorders may shed more light on this issue. Kessler and collegues31 showed that the 12-months prevalence of any mental disorder was more or less stable over a 10-years period, i.e. 29.5% and 26.2% respectively. With respect to the Netherlands, the NEMESIS study32 showed similar outcomes: in contrast to the expectations of mental health professional the 12-months prevalence remained stable at about 17%33 in a similar period. These important results suggest that, although the studies were conducted among the general population, it is more likely that the prevalence of assessed mental health problems was relatively stable and did not differ significantly across occupational groups over time.

31 Kessler RC, Chiu WT, Demler O, Walters EE (2005b). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Archives of General Psychiatry 62: 617-627.

32 Vollebergh WAM, Iedema J, Bijl RV, Graaf R de, Smit F, Ormel J (2001). The structure and stability of common mental disorders: the Nemesis-study. Archives of General Psychiatry 58: 597-603.

33 Graaf R de, Have M ten, Dorsselaer S van. De psychische gezondheid van de Nederlandse bevolking NEMESIS-2: Opzet en eerste resultaten [The mental health of the Dutch population NEMISIS-2: Design and first results. In Dutch]. Utrecht, Trimbos, 2012.

Dr. Huang suggested “It would be optimal if some indicators, such as socioeconomic indexes, could be adopted to adjust potential bias from different sampling time”. It is unclear to us which indices could be used.

Other comments
1. Abstract:
Page 3, line 41:
RET should be full-named when the abbreviation is first showed in the paper.

Response
We have provided the full name

2. Samples and procedures:
Page 6, line 31-46:
The time when Group1 and Group2 studies were conducted have not indicated. In addition, the code for study groups should be consistent in the paper. Group 1/Group 2, Police 2002/Police 2004 (Page 10, line 22-31), and Police A/Police B (in Table 2 and Table 3) all mixed up in the manuscript.

Response
We added the years of the surveys of both police groups (2002, 2004)
Based on previous comments, we have revised all tables, and have given the study groups the same name throughout the manuscript. We have corrected inconsistent codes of the study groups.

3. Samples and procedures:
Page 7, line 10-17:
The numbers are not equal (1116+149≠1294).

Response
We have corrected the numbers and controlled for other in-equalities throughout the manuscript. Since we had to re-analyze our data (see above), we have made the decision to exclude respondents with one or more missing values (listwise deletion). As a consequence all numbers are clearly the same in the manuscript and tables. We have explained this in the revised method section.

4. Samples and procedures:
Page 7, line 41:
The numbers for group 7 and group 8 in methods (281 and 238 respectively) are not consistent with the numbers in abstract for these groups (284 and 239) (Page 3, line 43).

Response
See our response to point 3 (previous point)

5. Measures:
Page 8, line 34:
For which year were the Dutch SCL-90-R cut-off scores for a normal population adopted in the study?

The cut-off scores should be an issue particularly because the data of different studies were collected at different periods.

Response
We used the norm tables from 1986 since these norm tables are more or less the standard and take into account differences between men and women. In the revised manuscript we mention the year 1986 explicitly.

Based on this comment and reviewers previous comments (general impression), we have added a section on this topic in the discussion part (see above).

6. Measures:
Page 8, line 46:
How did authors detect the clinical and sub-clinical levels for a brief scale of the SCL-90-R?

Response
The Dutch SCL-90-R norm tables make an explicit distinction between severe symptoms (80the percentile) and very severe symptoms (95th percentile). Very severe symptoms are considered to be clinically relevant. For example, according to the same Dutch norm tables the mean scores of a norm group of psychiatric patients on the subscales anxiety and depression were 26.0 (sd=9.9) and 41.9 (sd=14.8) respectively. In our total sample the mean scores of those with very severe anxiety and depression symptoms were 27.7 (sd=5.2) and 43.3 (sd=8.1) respectively.

Since other readers may have similar questions, we have added the following in the results section (pag 12):

Very severe symptoms such as very severe symptoms of depression are considered to be clinically relevant. According to the Dutch norm tables the mean scores of a norm group of psychiatric patients on the subscales anxiety and depression were 26.0 (sd=9.9) and 41.9 (sd=14.8) respectively. In our total sample the mean scores of those with very severe anxiety and depression symptoms were 27.7 (sd=5.2) and 43.3 (sd=8.1) respectively.

7. Results:
Page 10, line 22-31:
Why did the authors make ANOVA for Group1, Group2 and employees of supermarkets? Employees of supermarkets were not Group 9 who administered the brief scale of the SCL-90.
Tables 2 and 3:
Why were the data of group 11 not shown in tables? Instead, the results of Group 9 appeared in Tables. Did any mistake occur?

Response

We conducted an ANOVA because the group of employees of supermarkets were, as described in the measures section, administered a brief scale consisting of 17 random items of the SCL-90-R. We have no norm tables for these 17 items, and therefore they were not included in the logistic regression analyses. For this reason this sub group is not included in tables 2 and 3 (but is included in table 1 about the demographics). We did not calculate the sum scores of these 17 items among the other study groups because we believe these findings would be rather redundant, especially because we

examined both severe symptoms and very severe symptoms.

VERSION 2 – REVIEW

REVIEWER	Torbjørn Moum Professor emeritus Dept. of Behavioural Sciences in Medicine University of Oslo Norway No conflict of interest
REVIEW RETURNED	25-Nov-2012

GENERAL COMMENTS	Rates of psychological distress / mental health problems consistently are significantly higher among males than among females (at least when assessed multivariately). This pattern appears rather at odds with the gender differences commonly observed in broad population samples, and seems to warrant an explanatory note. Is this indicative of some peculiar kind of selection or is there a problem with the extremely skewed gender distribution for a couple of the occupations included (military and police).

REVIEWER	Shu-Ling Huang Associate Professor Department of Psychology Chung Shan Medical University, Taiwan No conflicts of interests
REVIEW RETURNED	30-Nov-2012

GENERAL COMMENTS	This study compared the mental health of police officers in Netherlands with those among other occupational groups. The result found that officers were not more at risk for serious mental health problems than any of the other occupational groups in this study. The findings are interesting and the discussion is appropriate. Only General Health Questionnaire (GHQ) and Impact of Event Scale (IES) on page 13 (line 20 and 29 respectively) should be full-named when the abbreviation is first showed in the paper.