A systematic review of the influence of occupational organophosphate pesticides exposure on neurological impairment

Noriko Takahashi, Masahiro Hashizume

**ABSTRACT**

**Objective:** The aim of this study was to conduct a systematic review of the published literature and to estimate whether or not there is a causal relationship between occupational exposure to organophosphate pesticides (OPs) and either neurological impairment or depressive symptoms.

**Data sources:** EMBASE, MEDLINE, Global Health and PsycINFO (1980 to April 2014).

**Setting:** Observational studies (cross-sectional, cohort and case–control studies) with exposed and unexposed groups.

**Participants:** People who occupationally use OPs for more than 1 month and their family.

**Primary outcome:** Results of neurological core test batteries or depressive symptoms such as headaches, anxiety and dizziness.

**Study appraisal and synthesis methods:** After an extensive search of various literature databases, one author screened titles and abstracts, searched the relevant publications manually and conducted data extraction. All extracted data from the selected articles were synthesised for analysis. Quality appraisal was conducted using the Newcastle Ottawa Scale.

**Results:** Of the 1024 articles retrieved by database search, 24 studies that met the inclusion and exclusion criteria were selected for analysis. Of the selected studies, 17 were cross-sectional and the remaining 7 were cohort and nested case–control studies. The geographical areas included in the studies were the USA (10 studies), the UK (4 studies), Africa (4 studies), Asia (3 studies), Europe (2 studies) and South America (1 study). Each of the included studies used different exposure and outcome assessments such as neurological scores and depressive symptoms, making it difficult to compare the results exactly. Most studies showed that exposed groups had poorer results than unexposed groups; however, owing to the inconsistent neurological test batteries, there was not enough pooling evidence to conduct a meta-analysis.

**Conclusions:** The findings of this literature review indicate that it is necessary to standardise the neurological or neuropsychological test battery and methods of measuring exposure to OPs.

**Strengths and limitations of this study**

- The article represents a systematic review of epidemiological studies on adverse effects on the human central nervous system by occupational organophosphate pesticides (OPs) exposure, with a quality appraisal of each study.
- The article identifies problematic issues of exposure and outcome assessments.
- Meta-analysis could not be applied because only a small number of pooled studies were available.
- In some studies, it was difficult to judge negative effects caused only by OPs, because mixed pesticides were used.

**INTRODUCTION**

Ever since organophosphate pesticides (OPs) were developed, they have been used to combat insects for public health purposes and to support agricultural productivity and manufacturing processes. Because pesticide ingestion is one of the leading suicide methods, a large number of epidemiological studies have investigated the relationship between high-level OPs exposure such as pesticide poisoning and accidents and acute health effects. It has been reported that high-level OPs exposure is significantly related to neurological or neuropsychological impairment. In contrast, few studies have reported associations between occupational or cumulative OPs exposure and negative effects on human health, although some research has examined the negative influence on young children of cumulative OPs exposure and others have investigated relationships between reproductive health and occupational OPs exposure. A high level of OPs exposure is known to have adverse effects on the human central nervous system (CNS); therefore, occupational or cumulative OPs exposure also has the potential to negatively affect the CNS.
However, very few epidemiological studies that have assessed the relationships between occupational OPs exposure and neurological or mental problems have been published. The objective of this systematic review is to verify whether or not occupational OPs exposure negatively affects the human CNS. To investigate this further, we summarised the epidemiological evidence for the relationship between occupational OPs exposure and mental and neuropsychological aggression, especially for occupational OP users, and some of the limitations associated with the various studies are discussed.

MATERIALS AND METHODS

Searching strategy for identification of published studies

We searched the published literature using the OvidSP search software to select relevant observational studies. A geographical restriction was not imposed; however, the search was restricted to studies published from 1980 to 2014. Population-based case–control studies were excluded from the systematic review because it was difficult to assess accurate exposure doses for these studies. Because various pesticides including OPs are currently easily available to everyone, it is highly likely that these pesticides have been obtained for personal use. For this reason, it is almost impossible to obtain past records of pesticide use by every individual. The literature search was limited to studies in humans and to reports published in English, and the review was limited to epidemiological studies. Moreover, unpublished studies and the grey literature (literature that has not been formally published) were not searched in this systematic review; therefore, we did not contact authors to find unpublished studies. Studies investigating OPs exposure through food and water contamination were also excluded. A search of the following four databases was carried out: EMBASE Classic + EMBASE (1980 to Week13 2014); Ovid MEDLINE (1980 to Week13 2014); Global Health (1980 to Week12 2014); and PsycINFO (1980 to Week14 2014).

A combination of free-text terms and explore terms was used to identify relevant articles. For exposure, the following search keywords were used: organophosphate*, organophosphorous, pesticide*, or insecticide* and organophosphate pesticide (explore map term). For outcome, the following search keywords were used: neuro*, psychiatri*, psycholog*, mental health, mental illness, mental disorder, or depressi*, depression (explore map term) and mental health (explore map term). For subjects, the following search keywords were used: occupation*, agricultu* or farm*. For study design, the following search keywords were used: epidemiolog*, cohort, cross-sectional, or case–control and epidemiology (explore map term). An initial systematic search in the titles and abstracts was conducted using a combination of all these search terms. A second manual search of the reference lists from the selected relevant articles was performed to explore or retrieve articles found in the initial search in order to find as many available studies as possible.

Criteria for selecting studies for review

Only original research articles meeting the inclusion and exclusion criteria described below were used in the final review.

**Inclusion criteria:**

1. Study design
2. Must be observational studies: cross-sectional, cohort and case–control studies.
3. Studies must have exposed and unexposed groups.
4. Participants
   A. The subjects in the exposed group must either use OPs occupationally, or there must be a probability of being exposed to OPs during their work.
   B. The families of occupational OP users can be treated as participants.
5. Exposure
   A. Participants must be exposed to OPs for at least 1 month.
   B. Seasonal workers who used OPs for more than 1 month must be included.
6. Outcome: Studies must have carried out some tests to assess damage to the CNS or have conducted a survey or an interview to identify depressive symptoms.
7. Exposure–outcome association: Results must be reported as some types of relative risks or mean scores.

**Exclusion criteria:**

1. Study design: Experimental and laboratory-based studies including animal studies were excluded.
2. Participants: Studies of mainly patients of pesticide poisoning were excluded.
3. Exposure: Studies that did not specify the type of pesticides were excluded.
4. Outcome: Studies examining damage to the peripheral nervous system due to OPs exposure were excluded.
5. Language: Studies published in a language other than English were excluded.

Definitions used for the review

**Definition of cumulative exposure**

A. People who used OPs in their jobs for at least 1 month and had the probability of inhaling ambient OPs and absorbing OPs by spraying and touching.

B. Families of OP users were included as subjects because they may have been exposed to OPs by washing clothes contaminated by OPs and/or by touching OP users.

**Definition of poor mental health**

A. Neurological or neuropsychological impairment
   a. People who had poorer results in neurological or neuropsychological test batteries than healthy people of the same age.
b. People who had short-memory loss; for example, people who had experienced memory loss for at least 3 months but less than 6 months.

B. Depressive symptom
a. People who, regardless of their age, had chronic depressive symptoms including headache, fatigue, dizziness, sleeplessness and eye problems.
b. People who were diagnosed with depression by clinical doctors.

Study selection process
Using the search terms listed above, a total of 1024 references were obtained: 515 from EMBASE Classic + EMBASE, 31 from PsycINFO, 196 from Global Health and 282 from Ovid MEDLINE (figure 1). However, 77 animal studies, 90 studies not in English studies and 12 studies that did not meet the time restrictions were excluded. Of the remaining 845 studies, 516 were excluded because of duplications. A manual search of the titles and abstracts of the remaining 329 references excluded a further 272 studies. The 21 remaining articles were fully reviewed, after which 12 studies were deemed to meet the inclusion and exclusion criteria.9–20 In addition, 12 articles identified by the manual search were added to the systematic review (figure 1). To include as many relevant studies as possible, studies published before 1980 that were found by the manual search were included to the list for review. Finally, these 24 studies were selected for data extraction.9–32

Data extraction, synthesis and analysis
Data extraction forms were created to compare relevant data collected from each of the 24 studies. The following data were extracted to assess heterogeneity of the included studies: title, authors, year published, number of participants in the exposed and unexposed groups, occupation, and demographic information such as mean age, sex, smoking status and geographical area. In addition, the following data were extracted to assess confounding factors and statistical models among the included studies: inclusion and exclusion criteria such as first language, alcohol consumption, injury experience, confounding factors and statistical methods used. The following data were extracted to assess exposure and outcomes: types of pesticides, exposure assessment and outcome assessment to measure the neurological or neuropsychological ability and results obtained. Tables containing the data that were obtained using the data extraction forms were constructed and analysed. p Values and 95% CIs were elicited from the articles to judge statistical uncertainty. When a study had investigated depressive symptoms, the information was collected and a table was constructed. The impact and statistical magnitude of depressive symptoms were represented using plus or minus signs ‘++’, ‘+’ and ‘−’, based on the p value or 95% CI of the studies. All data extraction, coding and quality appraisal were conducted only by the first author; therefore, no disagreement events occurred.

Figure 1 Flow diagram of search and review process (OP, organophosphates).
Quality appraisal
The quality of the 24 studies was appraised using a scale adapted from the ‘Newcastle/Ottawa Scale (NOS)’ (the appraisal standard of NOS is shown in online supplementary appendix A). Based on the NOS, each study was evaluated using the point system. When a study included relevant information that could be associated with the NOS, one point was added. Five items in cross-sectional studies and eight items in cohort and case-control studies that could be related to the NOS were identified. Therefore, cross-sectional studies assigned 5, 4, 3 or 0–2 points were evaluated as very good, good, satisfactory or unsatisfactory studies, respectively. Similarly, cohort/case-control studies with 7–8, 5–6, 4 and 0–3 points were identified as very good, good, satisfactory or unsatisfactory, respectively.

RESULTS
As a result of the search strategy described in the Materials and methods section, 12 studies were identified from the database search and another 12 studies were found after a manual search. These 24 articles, published between 1975 and 2010, met all the inclusion and exclusion criteria. A summary of the characteristics of the 24 selected articles is shown in Table 1.

Study design and geographical area
Of the selected studies, 17 were cross-sectional and the remaining 7 were cohort and nested case-control studies. The geographical areas included in the studies were the USA (10 studies), the UK (4 studies), Africa (4 studies; two in South Africa, 1 in Egypt and 1 in Kenya), Asia (3 studies; 2 in India and 1 in Sri Lanka), Europe (2 studies; 1 in Spain and 1 in Poland) and South America (1 study in Ecuador).

Characteristics of participants
Because the participants were limited to people who had the probability of being occupationally exposed to OPs, the majority of the participants (60–70%) were men. Most of the time, agricultural work such as pesticide application and farming is performed predominantly by men. Six of the 24 studies included male and female participants,9 11 17 25 27 32 and only 1 study used all female participants in the exposed and control groups.21 In 13 of the studies, the mean age of the exposed participants was in the 30s, in six studies the mean age was in the 40s,9 14 15 19 20 31 and in two studies the mean age was in the 50s.13 17 The mean age in two studies was 29, very close to 30.25 32 One of the studies did not report detailed demographic data of the participants.10

Source of recruitment and sample size
Ten of the 24 studies were on pesticide applicators including private, commercial, and tree, fruit and vegetable applicators. Five and three studies were on farmers and sheep farmers, respectively, and two studies were on factory workers and greenhouse workers. One study investigated depressive symptoms in the spouses of OPs users. In the study by Korsak et al.,22 the specific occupation of the population in the study was not stated; however, the participants had experienced occupational OPs exposure. The number of participants in the exposed groups varied from 16 to 2051, while the control groups had a wider range of participants (16–27 023).

Exposure assessment
Exposure assessment in the included studies could be divided, for the most part, into five patterns: indirect assessment using, for example, an interview or questionnaire; direct assessment including a measurement of urinary metabolites and acetylcholinesterase (AChE) levels in the blood or a measurement of ambient OPs using a patch and a pump; a combination of direct and indirect methods; a combination of a biomarker and OPs exposure levels included in house dust; and a combination of biomarkers and ambient OP levels. Seven of the 24 studies used indirect methods, and six studies used blood AChE inhibition levels to measure AChE levels in the blood as an exposure indicator. Six studies used a combination of indirect methods and biomarkers, three studies used biomarkers and ambient OP levels and one study used a biomarker and house dust. The remaining study did not mention any exposure assessment methods. In all the studies that used urinary metabolites as exposure assessment, the results were presented as the sum of dialkylyphosphates (DAP; ie, the sum of six DAP metabolites: DMP (dimethylphosphate), DMTP (dimethyldithiophosphate), DMDTP (dimethylthiophosphate), DMP (diethylphosphate), DETP (diethyldithiophosphate) and DEDTP (diethyldithiophosphate)).25 30–32

Outcome measurements
Two different outcome measurements were used in the studies; one measured neurological impairment and the other assessed depressive symptoms. Of the 24 studies, 19 used cognitive function tests to investigate negative neurological influences caused by OPs exposure.

Associations between outcome and exposure
Ten of the 19 studies that investigated cognitive impairment mentioned that at least one measure outcome showed more impairment in the exposed groups; however, these observations were not significant (p<0.05). Seven of the studies reported some significant positive associations of exposure with poor outcome (p<0.05); however, even in these cases, the significant decrements were observed only in some of the neurological tests, mainly in the Digit Span and Santa Ana Dexterity tests. Indeed, there are several versions of these neurological tests and the significance of the scores often depended on the versions of the tests that
Table 1: Findings reported in epidemiological studies into occupational low-level OPs exposure and mental illness

<table>
<thead>
<tr>
<th>Author</th>
<th>Study design</th>
<th>Country</th>
<th>Exposed population (N)</th>
<th>Chemical</th>
<th>Exposed assessment</th>
<th>Comparison group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Albers et al</td>
<td>CO USA</td>
<td>Chemical workers (53)</td>
<td>OP</td>
<td>Industrial HR, AChE INH, DR</td>
<td>Similar workers, not exposed (60)</td>
<td></td>
</tr>
<tr>
<td>2 Bazylewicz-Walczak et al</td>
<td>CO Poland</td>
<td>Greenhouse workers (26)</td>
<td>OP</td>
<td>QU</td>
<td>Greenhouse workers, not exposed (25)</td>
<td></td>
</tr>
<tr>
<td>3 Beseler et al</td>
<td>NC/CO USA</td>
<td>Case†: spouses of private applicators with depressive diagnoses (2051)</td>
<td>OP</td>
<td>QU or IN</td>
<td>Control: spouses of private applicators without depressive diagnoses (27 023)</td>
<td></td>
</tr>
<tr>
<td>4 Cole et al</td>
<td>CR Ecuador</td>
<td>Farmers, some applicators (144)</td>
<td>OP, CAR, FNG</td>
<td>IN, QU, AChE INH</td>
<td>Local population(72)</td>
<td></td>
</tr>
<tr>
<td>5 Daniell et al</td>
<td>CO USA</td>
<td>Farm worker applicators (49)</td>
<td>OP</td>
<td>QU, AChE INH</td>
<td>Slaughterhouse workers(40)</td>
<td></td>
</tr>
<tr>
<td>6 Dassanayake et al</td>
<td>CR Sri Lanka</td>
<td>Vegetable farm workers (38)</td>
<td>OP</td>
<td>NA</td>
<td>Hospital labourers (35)</td>
<td></td>
</tr>
<tr>
<td>7 Farahat et al</td>
<td>CR Egypt</td>
<td>Farm workers (52)</td>
<td>OP</td>
<td>AChE INH</td>
<td>Local population(50)</td>
<td></td>
</tr>
<tr>
<td>8 Fiedler et al</td>
<td>CR USA</td>
<td>Tree fruit farmers (57)</td>
<td>OP</td>
<td>QU, lifetime exposure metric</td>
<td>Cranberry/blueberry growers(low exposed), hardware storeowners (unexposed) (42)</td>
<td></td>
</tr>
<tr>
<td>9 Korsak et al</td>
<td>CR USA</td>
<td>Occupational exposure (16)</td>
<td>OP, CAR, OC</td>
<td>AChE INH</td>
<td>Local population (low exposure; 16)</td>
<td></td>
</tr>
<tr>
<td>10 Levin et al</td>
<td>CR USA</td>
<td>Pesticide applicators (24)</td>
<td>OP</td>
<td>IN, AChE INH</td>
<td>Farmers (24)</td>
<td></td>
</tr>
<tr>
<td>11 London et al</td>
<td>CR South Africa</td>
<td>Fruit farm pesticide applicators (163)</td>
<td>OP</td>
<td>QU(job-matrix)</td>
<td>Farm workers, not applicators (84)</td>
<td></td>
</tr>
<tr>
<td>12 London et al</td>
<td>CR South Africa</td>
<td>Fruit farm pesticide applicators (164)</td>
<td>OP</td>
<td>QU</td>
<td>Farm workers, not applicators (83)</td>
<td></td>
</tr>
<tr>
<td>13 Maizlish et al</td>
<td>CR USA</td>
<td>Pesticide applicators (46)</td>
<td>OP</td>
<td>UM, DR</td>
<td>Non-applicants (56)</td>
<td></td>
</tr>
<tr>
<td>14 Misra et al</td>
<td>PR India</td>
<td>Pesticide applicators (22)</td>
<td>OP</td>
<td>AChE INH</td>
<td>Hospital labourers (20)</td>
<td></td>
</tr>
<tr>
<td>15 Ohayo-Mitoko et al</td>
<td>CO Kenya</td>
<td>Farm worker applicators (256)</td>
<td>OP, CAR</td>
<td>AChE INH</td>
<td>Farm workers (low exposure; 152)</td>
<td></td>
</tr>
<tr>
<td>16 Rodnitzky et al</td>
<td>CR USA</td>
<td>Pesticide applicators (23)</td>
<td>OP</td>
<td>AChE INH</td>
<td>Farmers (23)</td>
<td></td>
</tr>
<tr>
<td>17 Roldan-Tapia et al</td>
<td>CR Spain</td>
<td>Greenhouse workers (40)</td>
<td>OP, CAR</td>
<td>QU, AChE INH</td>
<td>Local population (26)</td>
<td></td>
</tr>
<tr>
<td>18 Ross et al</td>
<td>CO UK</td>
<td>Sheep farmers (127)</td>
<td>OP</td>
<td>IN</td>
<td>Police workers (78)</td>
<td></td>
</tr>
<tr>
<td>19 Rothlein et al</td>
<td>CR USA</td>
<td>Farm workers (96)</td>
<td>OP</td>
<td>UM, House dust</td>
<td>Workers in hotels and tourist industry (45)</td>
<td></td>
</tr>
<tr>
<td>20 Srivastava et al</td>
<td>CR India</td>
<td>Manufacture workers (59)</td>
<td>OP</td>
<td>AChE INH</td>
<td>Manufacture workers, not exposed (17)</td>
<td></td>
</tr>
<tr>
<td>21 Steenland et al</td>
<td>CR USA</td>
<td>Termiticide applicators (191)</td>
<td>OP</td>
<td>IN,UM</td>
<td>Friends, blue collar workers (189)</td>
<td></td>
</tr>
<tr>
<td>22 Stephens et al</td>
<td>CR UK</td>
<td>Sheep farmers (146)</td>
<td>OP</td>
<td>QU</td>
<td>Quarry workers (143)</td>
<td></td>
</tr>
<tr>
<td>23 Stephens et al</td>
<td>CR UK</td>
<td>Sheep farmers (77)</td>
<td>OP</td>
<td>QU, UM</td>
<td>Quarry workers (69)</td>
<td></td>
</tr>
<tr>
<td>24 Stephens et al</td>
<td>CR UK</td>
<td>Orchard applicators (37)</td>
<td>OP</td>
<td>QU</td>
<td>Construction workers, pig farmers (57)</td>
<td></td>
</tr>
</tbody>
</table>

*Studies that included depressive symptoms for outcome assessments.
†Cases were defined as female spouses of private applicators who responded with ‘yes’ to the question ‘Has a DOCTOR ever told you that you had been diagnosed with depression requiring medication?’ Controls were female spouses who responded with ‘no’.
AChE, acetylcholinesterase; AChE INH, AChE inhibition; CAR, carbamates; CO, cohort; CR, cross-sectional; DR, dermal and respiratory absorption; FUN, fungicides; HR, hygiene records; IN, interview; NA, not applicable; NC, nested case–control; OC, organochlorines; OP, organophosphates; PR, prospective study; QU, Questionnaire; UM, urinary metabolites.
were used. Five studies used the Wechsler Adult Intelligence Scale (WAIS or WAIS-R), 25 26 four studies used the Neurobehavioral Evaluation System (NES), 27 two studies used the WHO Neurobehavioral Core Test Battery (NCTB) 28 29 and the remaining eight studies used their own scales.

Five studies adopted depressive symptoms as outcome measurements, as shown in table 2 ; however, the symptoms used in the studies were not standardised.

Statistical analysis
Sixteen studies used logistic regression, and the remaining eight used other statistical tests including $\chi^2$ test and t-test. Only one study adjusted for sex in the logistic regression. Fourteen of the 24 studies adjusted for age, and 12 of the 24 studies adjusted for education in the statistical analysis. However, only five studies adjusted for alcohol consumption before carrying out the statistical analysis, and only two studies adjusted for first language.

Methodological quality appraisal
Based on NOS, 5 of the 24 studies were of very good quality, 10 were of good quality and the remaining 9 were either satisfactory or unsatisfactory. Most studies with unsatisfactory scores either were carried out before 1990 or were performed in some of the less developed countries. In particular, the methods of recruitment of subjects, controlling for confounders, and outcome assessment were not appropriate. For example, in some studies, all of the participants were volunteers 14 28 and in another study, the subjects were not representative of the community from which they were recruited (factory workers). 20 In addition, in the unsatisfactory studies, how the outcome was assessed was not described, and methods needed to avoid confounders such as stratification and regression were not used. None of the cohort studies were assessed as being of very good quality because most of them did not have a long enough follow-up duration (in five studies, the duration was less than 6 months) and the selected subjects were not fully representative of the target community. Moreover, the methods of outcome assessment were not described in most of the cohort studies.

Data synthesis
The results of the neurological tests used in the studies are summarised in table 3 . As can be seen, the test batteries differed from study to study. The commonly used test batteries in NCTB, NES and WAIS were Symbol-Digit and Digit Span Forward and Backward. However, some studies that adopted NES and WAIS to measure neurological impairment implemented only a few subsets in the trials. Among the 13 studies that used a Symbol-Digit test, 4 used NES and unknown tests, 2 used WAIS and WAIS-R, and 1 used a Polish NCTB. Among the studies that used Digit Span Forward and Backward tests, some studies performed both tests, while the others performed only one of the tests as shown in table 3 . Overall, only four of the studies used the same test battery in NES and WAIS. Although three studies apparently used the same scoring systems, one of the scores was completely different from the scores in the other two studies. For example, in two studies, the scores reported by Daniell et al and Stephens et al were much lower and between 2.23 and 3.55. 12 20 Similarly, the mean scores reported by Bzylewicz-Walczak et al 21 were higher, 45.50 and 49.40, while the mean scores reported in the other studies were smaller, 22.8 and 22.3, in WAIS. 25 In consideration of the insufficient number of studies and possible systematic differences in the population characteristics and/or the measurement procedures between the studies, we decided not to conduct a meta-analysis.

DISCUSSION
The systematic keyword and manual searches of the published literature identified 24 epidemiological studies that examined the relationship between OPs and CNS.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Results obtained</th>
<th>Impact of outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beseler et al 10</td>
<td>Depression due to doctor’s diagnosis was not significantly related to low (OR 1.09; 95% CI 0.91 to 1.31) cumulative exposure</td>
<td>–</td>
</tr>
<tr>
<td>Levin et al 23</td>
<td>Anxiety score of the pesticide applicators was significantly higher (p&lt;0.05) than that of the farmers. However, there was no significant difference in measures of depression</td>
<td>++</td>
</tr>
<tr>
<td>London et al 24</td>
<td>Dizziness, sleepiness and headache had a significantly higher overall neurological symptom score (p&lt;0.05)</td>
<td>++</td>
</tr>
<tr>
<td>Misra et al 26</td>
<td>Common symptoms were headache (59%), giddiness (50%), ocular symptoms (27%) and paraesthesia (18%), and no neurological change was seen</td>
<td>–</td>
</tr>
<tr>
<td>Ohayo-Mitoko et al 27</td>
<td>A significant change in symptom prevalence was found for the respiratory (2.48% CI (0.78% to 5.38%) and central nervous system (2.56% CI (0.99% to 6.62%), but in terms of skin symptoms and eye symptoms, there was no statistically significant change</td>
<td>++</td>
</tr>
</tbody>
</table>
When the relevant information was assessed, two main findings were obtained; one was the method of exposure assessment, and the other was the method used for the outcome measurement. For exposure assessment, the measurement methods were categorised as direct, indirect and a combination of direct and indirect. For the outcome measurements, two main assessments were used: neurological impairment and depressive symptoms.

**Exposure assessment**

Exposure assessment was not used for group allocation in all the studies; rather, it was implemented to measure how many participants were exposed and the outcomes of the neurobehavioural tests. Different exposure assessment methods were used in each study, which made it difficult to accurately compare the studies. In addition, there seemed to be methodological imperfections in the direct and indirect methods. For example, in one study, an interview and questionnaire were used for recruiting participants above 60 years of age who had been retired for 11 years.17 This method is subject to recall bias because the rate of cognitive impairment is likely to have increased as the participants aged. However, other indirect methods, especially extensive history records of pesticide use, could be considered as a proxy of how much OPs might have accumulated in the body; thus, records of this type can be used to estimate the amount of OPs by long-term exposure, even though there may be some recall bias. For the direct methods, DPA or urinary metabolite was used as an exposure index in the study; however, DPA is metabolised rapidly and excreted.6 In contrast, blood AChE levels take approximately 1 week to become normal;39 hence, although blood AChE levels cannot be used to assess the accumulation of OPs in body tissues over a long time, it can be used to assess short-term exposure. To minimise measurement errors, a mixed method for the assessment of short-term and long-term exposure should be established.

**Outcome assessment**

The main problem in analysing the outcome measurements was the inconsistencies in neurological test batteries. Various versions of the neurological tests were used in the studies and the content of the tests differs slightly in each study (table 3). Therefore, only a few tests were common across some of the studies, which made it difficult to compare the studies. Further, a meta-analysis could not be applied because of the insufficient number of studies. Meta-analysis could have been performed by dividing the results into subgroups; however, the results could be highly misleading because of loss of power.40 In terms of depressive symptoms, the outcome assessment was again different in each study. For instance, one study used the proportion of headaches, while another used dizziness and sleepiness as the main outcomes. To gain better insights into whether occupational OP exposure can negatively affect the human CNS, at the very least, neurological test batteries should be standardised.
and guidelines for measuring of neurological symptoms should be set for all future epidemiological studies. Furthermore, although some studies mentioned the possible relationship between OPs exposure and confounding factors such as age and education, statistical tests between the exposed and control groups were not performed in these studies. These inconsistencies make it difficult to compare the neurological impairment outcomes among the studies.

**Study design**

Although 17 of 24 studies were cross-sectional studies, longitudinal or cohort studies are more appropriate because agricultural work using pesticides is easily influenced by seasonality. One research regarding reproductive health by OPs exposure stated that sperm concentration and counts are negatively affected in spring, the peak season, rather than winter. Therefore, the effect on the CNS could also be affected by seasonality.

**Sources of possible biases**

Since only published studies written in English were searched, publication bias could have occurred. In future studies, non-English studies and unpublished studies should be included to reduce publication bias. In trials that included foreign workers, first language and education levels could be considered as possible biases because there is a possibility that non-native participants did not fully understand the content and instructions for the tests, which could lead to their obtaining a lower score than native speakers. Additionally, the education systems in developed and less developed countries could be very different. Nowadays, developed countries such as the USA and the Gulf countries have accepted foreign workers as an important part of the workforce. These factors needed to be adjusted carefully in the sampling and analytical stages of the study; however, only two of the selected studies mentioned first language in their statistical analyses. Occupation could also contribute to selection bias because, for example, a police officer or a construction worker would have a higher probability of experiencing loss of consciousness due to accidents than workers with different occupations.

**Possible confounders**

Apart from common confounders such as age and education, head injury and alcohol consumption could be other confounders, because they can cause neurological impairment due to memory deterioration. Although some of the studies adjusted for alcohol consumption in the analysis, no study adjusted for head injury. Furthermore, nutrition status including vitamin deficiency can also be relevant to the outcome of neurological tests. Thus, factors other than the common confounders that could negatively affect cognitive function should be adjusted for in the analysis.

**Strengths and limitations of this review**

A major strength of this systematic review is that the characteristics of the selected studies were summarised using tables, and limitations of the exposure and outcome assessments used in these studies were identified mainly on the basis of the constructed tables. Furthermore, the systematic review allowed us to propose recommendations that will be useful for standardising future epidemiological research.

All of the selected studies were relevant to occupational OPs exposure; however, some of them included other pesticides such as carbamates, fungicides and herbicides. Pesticides that are commonly used in agriculture are usually mixtures of different pesticides, which are used to increase their effect. Four of the 24 selected studies used a combination of OPs, organochlorines, carbamates and fungicide; hence, the effect of only occupational OPs exposure could not be measured in these studies. In the outcome assessments, different neurological types of tests were used; consequently, the lack of pooling evidence meant that a meta-analysis could not be performed. Furthermore, the exclusion of studies written in languages other than English is another limitation of this review, and literature retrieval by only the first author could have introduced some bias into the selection of the studies.

**CONCLUSION**

The items tested in the neurological or neuropsychological test batteries, and the estimates of OPs exposure were inconsistent because they depended on the preferences of the investigators. For future studies, it would be best to standardise the neurological and neuropsychological test types, test batteries and the methods used to measure OPs, to enable precise comparisons of results and pooling of evidence from a large number of studies for future analyses. However, this may be difficult to achieve in practice because OPs are used in differing settings around the world, and education systems vary considerably between countries.

**Acknowledgements**

The authors thank Professor Ben Armstrong for his insightful comments on our paper.

**Contributors**

NT conceived the study design and participated in protocol development, literature searching, data extraction and data analysis and also drafted the manuscript. MH critically reviewed the draft and contributed to the manuscript revisions.

**Funding**

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests**

None.

**Provenance and peer review**

Not commissioned, externally peer reviewed.

**Data sharing statement**

No additional data are available.

**Open Access**

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/
REFERENCES


35. Anger WK. Neurobehavioral tests and systems to assess neurotoxic exposures in the workplace and community. Occup Environ Med 2003;60:531–8, 474.


Appendix A
The Appraisal Standard of Newcastle/Ottawa Scale

Selection

1) Representativeness of the exposed group/cohoot
   a) Truly representative of the average farmers or pesticides applicators in the community*
   b) Somewhat representative of the average farmers or pesticides applicators in the community*
   c) Selected group of users (e.g. factory workers, volunteers)
   d) No description of the derivation of the group

2) Selection of the non-exposed group/cohoot
   a) Drawn from the same community as the exposed group*
   b) Drawn from a different source
   c) No description of the derivation of the non-exposed group

3) Ascertainment of exposure
   a) Secure record (e.g. biomarkers)*
   b) Structured interview or questionnaire*
   c) Written self reports
   d) No description

4) Demonstration that outcome of interest was not present at start of study (Cohort Studies Only)
   a) Yes*
   b) No

Confounder

1) Comparability of groups on the basis of the design or analysis
   a) Study controls for age and education*
   b) Study controls for any additional factor* (e.g. alcohol consumption, smoking, and first language)

Outcome

1) Assessment of outcome
   a) Independent blind assessment*
   b) Record linkage*
   c) Self reports
d) No description

2) Was follow-up long enough for outcomes to occur (Cohort Studies Only)
   a) Yes (select an adequate follow up period for outcome of interest)*
   b) No

3) Adequacy of follow up of cohorts (Cohort Studies Only)
   a) Complete follow up – all subjects accounted for*
   b) Subjects lost to follow up unlikely to introduce bias – small number lost - > 70% follow up, or description provided of those lost*
   c) Follow up < 70% and no description of those lost
d) No statement

Case Control Studies:

Selection

1) Is the case definition adequate?
   a) Yes, with independent validation*
   b) Yes, e.g. record linkage on self reports
   c) No description

2) Representativeness of the cases
   a) Consecutive or obviously representative series of cases*
   b) Potential for selection biases or non stated

3) Selection of Controls
   a) Community controls*
   b) Hospital controls
   c) No description

4) Definition of Controls
   a) No history of disease (endpoint)*
   b) No description of source

Confounder

1) Comparability of cases and controls on the basis of design or analysis
   a) Study controls for age and education*
   b) Study controls for any additional factor*
Exposure

1) Ascertainment of exposure
   a) Secure record (e.g. biomarkers)*
   b) Structured interview where blind to case/control status*
   c) Interview not blinded to case/control status
   d) Written self reports or medical record only
   e) No description

2) Same method of ascertainment for cases and controls
   a) Yes*
   b) No

3) Non-Response rate
   a) Same rate for both groups*
   b) Non respondents described
   c) Rate different and no designation

*: plus one point

There are five items in cross-sectional studies and eight items in cohort and case control studies, respectively. The quality of the studies was defined as follows.

Cross-sectional Studies:
Very Good Studies: 5 points
Good Studies: 4 points
Satisfactory Studies: 3 points
Unsatisfactory Studies: 0 to 2 points

Cohort / Case control Studies:
Very Good Studies: 7 to 8 points
Good Studies: 5 to 6 points
Satisfactory: 4 points
Unsatisfactory Studies: 0 to 3 points
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Representativeness of the exposed group</td>
<td>a) (+1)</td>
<td>b) (+1)</td>
<td>c) (0)</td>
<td>a) (+1)</td>
<td>b) (+1)</td>
<td>b) (+1)</td>
</tr>
<tr>
<td>a) Truly representative of the average farmers or pesticides applicators in the community</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) Somewhat representative of the average or pesticides applicators in the community</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) Selected group of users</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) No description of the derivation of the group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2) Selection of the non exposed group</td>
<td>a) (+1)</td>
<td>b) (0)</td>
<td>b) (0)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
<td>b) (0)</td>
</tr>
<tr>
<td>a) Drawn from the same community as the exposed group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) Drawn from a different source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) No description of the derivation of the non exposed group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3) Ascertainment of exposure</td>
<td>b) (+1)</td>
<td>d) (0)</td>
<td>a) (+1)</td>
<td>b) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td>a) Secured record (e.g. biomarkers)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) Structured interview or questionnaire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C) Written self report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) No description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confounders</td>
<td>b) (+1)</td>
<td>- (0)</td>
<td>a) (+1)</td>
<td>- (0)</td>
<td>- (0)</td>
<td>- (0)</td>
</tr>
<tr>
<td>1) Comparability of groups on the basis of the design or analysis</td>
<td>b) (+1)</td>
<td></td>
<td>a) (+1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) Study controls for age and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
education
b) Study controls for any additional factor (e.g. alcohol consumption, smoking, and first language)

### Outcome
1) Assessment of outcome

<table>
<thead>
<tr>
<th></th>
<th>a) Independent blind assessment</th>
<th>b) Record linkage</th>
<th>c) Self report</th>
<th>d) No description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) (+1)</td>
<td>b) (+1)</td>
<td>d) (0)</td>
<td>b) (+1)</td>
<td>d) (0)</td>
</tr>
<tr>
<td>a) (+1)</td>
<td>b) (+1)</td>
<td>d) (0)</td>
<td>b) (+1)</td>
<td>d) (0)</td>
</tr>
<tr>
<td>a) (+1)</td>
<td>b) (+1)</td>
<td>d) (0)</td>
<td>b) (+1)</td>
<td>d) (0)</td>
</tr>
</tbody>
</table>

**Overall Score**
- 5/5 Very Good
- 2/5 Unsatisfactory

2) Selection of the non exposed group

<table>
<thead>
<tr>
<th></th>
<th>a) Truly representative of the average farmers or pesticides applicators in the community</th>
<th>b) Somewhat representative of the average or pesticides applicators in the community</th>
<th>c) Selected group of users</th>
<th>d) No description of the derivation of the group</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) (+1)</td>
<td>b) (+1)</td>
<td>a) (+1)</td>
<td>c) (0)</td>
<td>c) (0)</td>
</tr>
<tr>
<td>a) (+1)</td>
<td>b) (+1)</td>
<td>a) (+1)</td>
<td>c) (0)</td>
<td>c) (0)</td>
</tr>
<tr>
<td>a) (+1)</td>
<td>b) (+1)</td>
<td>a) (+1)</td>
<td>c) (0)</td>
<td>c) (0)</td>
</tr>
</tbody>
</table>

### Table 1 Continued

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Representativeness of the exposed group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) Truly representative of the average farmers or pesticides applicators in the community</td>
<td>b) (+1)</td>
<td>a) (+1)</td>
<td>c) (0)</td>
<td>c) (0)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td>b) Somewhat representative of the average or pesticides applicators in the community</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) Selected group of users</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) No description of the derivation of the group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2) Selection of the non exposed group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) Drawn from the same community as the exposed group</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
<td>c) (0)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td>b) Drawn from a different source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) No description of the derivation of the non exposed group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 3) Ascertainment of exposure

<table>
<thead>
<tr>
<th>Method</th>
<th>a)</th>
<th>b)</th>
<th>c)</th>
<th>d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Secured record (e.g. biomarkers)</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(+1)</td>
</tr>
<tr>
<td>b) Structured interview or questionnaire</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(+1)</td>
</tr>
<tr>
<td>C) Written self report</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) No description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Confounder

1) Comparability of groups on the basis of the design or analysis

<table>
<thead>
<tr>
<th>Method</th>
<th>a)</th>
<th>b)</th>
<th>c)</th>
<th>d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Study controls for age and education</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(-0)</td>
</tr>
<tr>
<td>b) Study controls for any additional factor (e.g. alcohol consumption, smoking, and first language)</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(-0)</td>
</tr>
</tbody>
</table>

### Outcome

1) Assessment of outcome

<table>
<thead>
<tr>
<th>Method</th>
<th>a)</th>
<th>b)</th>
<th>c)</th>
<th>d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Independent blind assessment</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(+1)</td>
</tr>
<tr>
<td>b) Record linkage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) Self report</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) No description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Overall Score

<table>
<thead>
<tr>
<th>Source</th>
<th>1)</th>
<th>2)</th>
<th>3)</th>
<th>4)</th>
<th>Overall Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rothlein et al 2006</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(+1)</td>
<td>5/5 Very Good</td>
</tr>
<tr>
<td>Srivastava et al 2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4/5 Good</td>
</tr>
<tr>
<td>Steenland et al 2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4/5 Good</td>
</tr>
<tr>
<td>Stephens et al 1995</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1/5 Unsatisfactory</td>
</tr>
<tr>
<td>Stephens et al 1996</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stephens et al 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5/5 Very Good</td>
</tr>
</tbody>
</table>

Continued…

### Selection

<table>
<thead>
<tr>
<th>Method</th>
<th>a)</th>
<th>b)</th>
<th>c)</th>
<th>d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Representativeness of the exposed group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) Truly representative of the average farmers or pesticides applicators in the community</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(+1)</td>
<td>(+1)</td>
</tr>
<tr>
<td>b) Somewhat representative of the average or pesticides applicators in the community</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) Selected group of users</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) No description of the derivation of the group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2) Selection of the non exposed group</td>
<td>a) Drawn from the same community as the exposed group</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>b) Drawn from a different source</td>
<td></td>
<td>a) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td>c) No description of the derivation of the non exposed group</td>
<td></td>
<td>a) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td>3) Ascertainment of exposure</td>
<td>a) Secured record (e.g. biomarkers)</td>
<td>b) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td></td>
<td>b) Structured interview or questionnaire</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C) Written self report</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>d) No description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Confounder</strong></td>
<td>1) Comparability of groups on the basis of the design or analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>a) Study controls for age and education</td>
<td>a) (+1)</td>
<td>- (0)</td>
<td>b) (+1)</td>
</tr>
<tr>
<td></td>
<td>b) Study controls for any additional factor (e.g. alcohol consumption, smoking, and first language)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
<td>1) Assessment of outcome</td>
<td>b) (+1)</td>
<td>d) (0)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td></td>
<td>a) Independent blind assessment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b) Record linkage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c) Self report</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>d) No description</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Overall Score                   | 5/5 Very good                                   | 2/5 Unsatisfactory | 5/5 Very Good | 4/5 Good | 4/5 Good | 4/5 Good |
## Table 2: Quality Appraisal (Cohort Studies)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Representativeness of the exposed cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) Truly representative of the average farmers or pesticides applicators in the community</td>
<td>c) (0)</td>
<td>c) (0)</td>
<td>a) (+1)</td>
<td>b) (+1)</td>
<td>c) (0)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td>b) Somewhat representative of the average or pesticides applicators in the community</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) Selected group of users</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) No description of the derivation of the cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2) Selection of the non exposed cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) Drawn from the same community as the exposed cohort</td>
<td>b) (0)</td>
<td>a) (+1)</td>
<td>b) (0)</td>
<td>a) (+1)</td>
<td>b) (0)</td>
<td>b) (0)</td>
</tr>
<tr>
<td>b) Drawn from a different source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) No description of the derivation of the non exposed cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3) Ascertainment of exposure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) Secured record (e.g. biomarkers)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
<td>b) (+1)</td>
<td>a) (+1)</td>
<td>b) (+1)</td>
</tr>
<tr>
<td>b) Structured interview or questionnaire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C) Written self report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) No description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4) Demonstration that outcome of interest was not present at start of study</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td>a) Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Confounders</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1) Comparability of groups on the basis of the design or analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) Study controls for age and education</td>
<td>- (0)</td>
<td>a) (+1)</td>
<td>b) (+1)</td>
<td>- (0)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td>b) Study controls for any additional factor (e.g. alcohol consumption,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2 Continued

<table>
<thead>
<tr>
<th>Outcome</th>
<th>1) Assessment of outcome</th>
<th>2) Was follow-up long enough for outcomes to occur</th>
<th>3) Adequacy of follow up of cohorts</th>
<th>Overall Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a) Independent blind assessment</td>
<td>b) (+1)</td>
<td>d) (0)</td>
<td>d) (0)</td>
</tr>
<tr>
<td></td>
<td>b) Record linkage</td>
<td>d) (0)</td>
<td>d) (0)</td>
<td>d) (0)</td>
</tr>
<tr>
<td></td>
<td>c) Self report</td>
<td>d) (0)</td>
<td>d) (0)</td>
<td>d) (0)</td>
</tr>
<tr>
<td></td>
<td>d) No description</td>
<td>d) (0)</td>
<td>d) (0)</td>
<td>d) (0)</td>
</tr>
<tr>
<td></td>
<td>a) Yes (select adequate follow up period for outcome of interest)</td>
<td>b) (0)</td>
<td>b) (0)</td>
<td>b) (0)</td>
</tr>
<tr>
<td></td>
<td>b) No</td>
<td>d) (0)</td>
<td>d) (0)</td>
<td>d) (0)</td>
</tr>
<tr>
<td></td>
<td>a) Complete follow up-all subjects accounted for</td>
<td>b) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td></td>
<td>b) Subjects lost to follow up unlikely to introduce bias- small number lost- &gt;70% follow up, or description provided of those lost</td>
<td>b) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td></td>
<td>c) Follow up rate&lt;70% and no description of those lost</td>
<td>b) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
</tr>
<tr>
<td></td>
<td>d) No statement</td>
<td>b) (+1)</td>
<td>a) (+1)</td>
<td>a) (+1)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3 Quality Appraisal (Case-control Studies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection</td>
</tr>
<tr>
<td>1) Is the case definition adequate?</td>
</tr>
<tr>
<td>a) Yes, with independent validation</td>
</tr>
<tr>
<td>b) Yes, e.g. record linkage or based on self reports</td>
</tr>
<tr>
<td>C) No description</td>
</tr>
<tr>
<td>2) Representativeness of the cases</td>
</tr>
<tr>
<td>a) Consecutive or obviously representative series of cases</td>
</tr>
<tr>
<td>b) Potential for selection biases or not stated</td>
</tr>
<tr>
<td>3) Selection of Controls</td>
</tr>
<tr>
<td>a) Community controls</td>
</tr>
<tr>
<td>b) Hospital controls</td>
</tr>
<tr>
<td>C) No description</td>
</tr>
<tr>
<td>4) Definition of Controls</td>
</tr>
<tr>
<td>a) No history of disease (endpoint)</td>
</tr>
<tr>
<td>b) No description of source</td>
</tr>
<tr>
<td>Confounders</td>
</tr>
<tr>
<td>1) Comparability of cases and controls on the basis of design or analysis</td>
</tr>
<tr>
<td>a) Study control for age and education</td>
</tr>
<tr>
<td>b) Study controls for any additional factor</td>
</tr>
<tr>
<td>Exposure</td>
</tr>
<tr>
<td>1) Ascertainment of exposure</td>
</tr>
<tr>
<td>a) Secure record (biomarkers)</td>
</tr>
<tr>
<td>b) Structured interview where blind to case/control status</td>
</tr>
<tr>
<td>c) Interview not blinded to case/control status</td>
</tr>
<tr>
<td>d) Written self report or medical record only</td>
</tr>
<tr>
<td>2) Same method of ascertainment for cases and controls</td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>a) Yes</td>
</tr>
<tr>
<td>b) No</td>
</tr>
<tr>
<td>3) Non-response rate</td>
</tr>
<tr>
<td>a) Same rate for both groups</td>
</tr>
<tr>
<td>b) Non respondents described</td>
</tr>
<tr>
<td>c) Rate different and no designation</td>
</tr>
<tr>
<td>Overall Score</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Table 3 Continued