Review of occupational exposure to manganese and the potential health effects of such exposure

Sharea Ijaz

Jos Verbeek

Anneli Ojajärvi

Markku Sainio

Grant McMillan

Finnish Institute of Occupational Health

Cochrane Occupational Safety and Health Review Group

PO Box 310

FI 70101 Kuopio, Finland

Correspondence: jos.verbeek@ttl.fi

Tel +358 46 8108709

Contents

Summary

Manganese (Mn) is the fourth most widely used element. It occurs naturally in the earth and must be mined and processed before it can be used commercially. Manganese is an essential trace element for humans but its excess can also be toxic leading to disorders of the nervous system. Exposure at work exists to a variable extent in jobs where manganese is an essential part of the process or the equipment used. Examples are mining, milling ore, welding, steel manufacture and the production of dry batteries. The objective of the project was to summarise and assess the evidence on potential adverse health effects of manganese for workers, as well as to quantify the evidence in a exposure-dose-response relationship.

Methods: We searched multiple electronic databases including those of grey literature, and supplemented these searches with reference screening and contacting experts. We included systematic reviews and meta-analyses addressing occupational manganese exposure. We also included prospective and retrospective follow up studies that compared manganese exposed workers to unexposed workers, or those that compared workers exposed to high levels of manganese to workers exposed to low levels. We excluded cross-sectional studies. Any adverse health outcome was included. We did study selection, data extraction and risk of bias assessments duplicate and we agreed upon with discussion. We synthesized the evidence largely narratively due to a lack of data reported in the studies

Results: Thirteen systematic reviews or meta-analyses and eight primary studies were included.

We found no evidence for what constitutes a control group without exposure, conflicting evidence on the NOAEL (no observed adverse effect level) or the usefulness of estimating the concentration of Mn in blood or urine as measures of exposure.

For all health outcomes, the evidence was very low quality according to GRADE and insufficient for a causal association according to the Danish Occupational Medicine Association. We found uncertainty ranging from no risk (RR 0.91 95% Confidence Interval (CI) 0.80 to 1.04) to a very large risk of Parkinson's disease (RR 4.92 95% CI 0.96 to 25.22) due to Mn exposure. For neurological test outcomes, the exposure did not affect finger-tapping-scores considerably based on data from three studies (Mean Difference (MD) -0.05, 95% CI -0.24 to 0.14). For other neurological test outcomes, evidence was inconsistent. For respiratory disease and cancer, there was very low quality evidence from a single study each that Mn did not lead to a considerably increased risk for either type of diseases in the exposed. We found no studies on fertility problems in men that met the inclusion criteria.

Conclusion: In the included studies, there is insufficient evidence to establish a causal link between occupational manganese exposure and any adverse health effects. This does not mean the absence of an effect but it means that the available studies simply do not support nor refute a link between Manganese and adverse health outcomes. . Future primary studies need to better define exposure such as what represents a non-exposed control group and assessment such as an indication of the total Mn content and the respirable fraction, and should have a prospective follow-up. Consensus is needed on the most valid and reliable outcomes of interest for this question.

Resume

Mangan (Mn) er den fjerde mest udbredte element. Det forekommer naturligt i jorden og skal udvindes og behandles, før det kan anvendes kommercielt. Mangan er et vigtigt sporelement for mennesker, men dets overskud kan også være giftig og føre til lidelser i nervesystemet. Eksponering på arbejdspladsen findes i varierende grad i jobs, hvor mangan er en vigtig del af processen eller det anvendte udstyr. Eksempler er minedrift, fræsning af malm, svejsning, fremstilling af stål og fremstilling af tørbatterier. Formålet med projektet var at sammenfatte og vurdere beviserne på potentielle sundhedsskadelige virkninger af mangan for ansatte, samt at kvantificere bevismateriale i et eksponering-dosis-respons-forhold.

Metoder: Vi gennemsøgte flere elektroniske databaser herunder akademisk litteratur, og supplerede disse søgninger med henvisnings screening og ved at kontakte eksperter. Vi inkluderede systematiske bedømmelser og meta-analyser der omhandler erhvervsmæssig mangan eksponering. Vi inkluderede også prospektive og retrospektive opfølgende undersøgelser, der sammenlignede ansatte som blev udsat for mangan med ueksponerede ansatte, eller dem, der sammenlignede ansatte, der udsættes for høje niveauer af mangan til ansatte udsat for lave niveauer. Vi ekskluderede tværsnitsstudier. Enhver negativ indvirkning på helbredet blev inkluderet. Vi foretog studie udvælgelse, udtræk af data og bedømte risikoen for partiske vurderinger og dublering, som vi blev enige om efter diskussion. Vi sammenfattede beviserne i vid udstrækning narrativt på grund af mangel på data indberettet i studierne

Resultater: Tretten systematiske bedømmelser eller meta-analyser og otte primære studier blev inkluderet.

Vi fandt ingen beviser for, hvad der udgør en kontrolgruppe uden eksponering, modstridende beviser på NOAEL (No Observed Adverse Effect Level) eller anvendeligheden ved at estimere koncentrationen af Mn i blod eller urin som mål for eksponering.

Gældende for alle sundhedsresultater var beviserne ifølge GRADE af en meget lav kvalitet og utilstrækkelig til en årsagssammenhæng ifølge Dansk Selskab for Arbejds- og Miljømedicin. Vi fandt usikkerhed der spænder fra ingen risiko (RR 0,91, 95% konfidensinterval (CI) 0,80-1,04) til en meget stor risiko for Parkinsons (RR 4,92 95% CI 0,96-25,22) grundet Mn eksponering. For neurologiske test resultater, påvirkede eksponeringen ikke Finger Tapping resultaterne betydeligt, baseret på data fra tre studier (Mean Difference (MD) -0,05, 95% CI -0,24 til 0,14). Med hensyn til andre neurologiske test resultater, var beviserne inkonsekvente. Med hensyn til luftvejssygdomme og kræft, var der beviser af meget lav kvalitet fra én enkel studie, som fastslåede at Mn ikke førte til en væsentlig øget risiko for begge typer af sygdomme i det udsatte. Vi fandt ingen studier om fertilitetsproblemer hos mænd, der opfyldte inklusionskriterierne.

Konklusion: I de inkluderede studier er der ikke tilstrækkelige beviser til at fastslå en årsagssammenhæng mellem erhvervsmæssig mangan eksponering og eventuelle sundhedsskadelige virkninger. Dette betyder ikke fravær af en virkning, men det betyder, at de foreliggende undersøgelser simpelthen ikke understøtter eller afkræfter en sammenhæng mellem mangan og negative sundhedsresultater. I fremtidige primære studier er det nødvendigt bedre at

definere eksponeringen, såsom hvad repræsenterer en ikke-eksponeret gruppe og en vurdering såsom en indikation af det samlede Mn-indhold og den respirable del, og bør have en fremadrettet opfølgning. Konsensus er nødvendig for de mest valide og pålidelige resultater som er af interesse for dette spørgsmål.

Background

Manganese (Mn) occurs naturally in the earth. It is the fourth most widely used metal in the world. An essential trace element for humans, it serves as a co-factor for important metalloenzyme activities within human cells. These help the body to perform a broad spectrum of functions including formation of connective tissue and bones, clotting factors, lipid and carbohydrate metabolism, and are crucial in a range of reactions in the CNS, which help maintain normal brain and nerve activity. While Mn deficiency can lead to illnesses related to bone, joint and collagen functions, its excess can also be toxic leading to disorders of the nervous system as reported in toxicological and pathological studies. [3,9]

The adequate dietary intake for adult men and women is 2.3 and 1.8 mg/day, respectively. A Tolerable Upper Intake Level (UL) of 11 mg/day was set for adults based on a no-observed-adverse-effect level for Western diets. [1]

Exposure to manganese can occur from various sources. These include diet, where it occurs naturally in water, grains and green leafy vegetables, and air, where it reaches from both natural and manmade sources. The most important manmade sources associated to Mn are, among others, mining and milling of manganese, ferroalloy production, iron and steel foundries, welding, battery production and power plants. Mn is also found in some unleaded gasoline as methylcyclopentadienyl manganese tricarbonyl (MMT) and high traffic levels have been found associated with higher Mn levels in the air.[2] Furthermore, plant fertilizers often contain manganese along with other metals, and the making of pigments, dyes, inks, and incendiary devices also involves manganese. Long-term parenteral nutrition can also lead to manganese toxicity.[3] Manganese could thus enter the body via the enteral (oral ingestion), parenteral (injection), or inhalation route, although occupational relevance is for ingestion and inhalation routes alone. NIOSH lists the potentially affected organs to be the respiratory system, central nervous system, liver and kidneys.[4]

There is a safety margin between daily requirement levels and those causing damage. When ingested or inhaled in amounts much greater than the daily requirements manganese can also damage cells. The various mechanisms of action based on current evidence suggest that manganese in large amounts increases oxidative stress within cells. Oxidative stress is what happens when there are not enough antioxidants to neutralize free radicals. Free radicals are the unstable molecules that react with other substances in the body and can damage cells. This can cause mitochondrial dysfunction, glutamate mediated excito-toxity and aggregation of proteins in the cell.[5-7]

The major neurodegenerative effect associated to excessive exposure to manganese is manganism. The clinical and other manifestations of manganism and how it can be differentiated from other movement disorders, especially idiopathic Parkinson's disease, with which its shares similarities are well documented. However, to our knowledge, epidemiological research has never formally applied these clinical criteria. Confidence in clinical diagnosis is increased by MRI and PET scans and tests of response to dopaminomimetric drugs and EDTA chelation. Both manganism and Parkinson's disease show abnormal features and pathophysiology in the basal ganglia. The basal ganglia are the part of the brain that is associated with control of voluntary motor movements and manganese

accumulates there.[8] However, in Parkinson's disease the findings relate to dopaminergic dysfunction of the nigrostriatal system manifesting as FDOPA uptake in PET imaging and positive response to levodopa treatment, which lack or are less evident in manganism.[9] Possibly, because of declining occupational exposures, there are no recent reports of frank manganism. More subtle adverse nervous system effects are studied with very sensitive but non-specific psychometric tests and functional brain imaging. These are called subclinical because they are assumed to be present before overt disease symptoms occur. Researchers compare the findings with these sensitive methods usually against a control group. The findings may be within normative values (which are unknown for some of these tests).

Besides neuronal damage, extensive exposure to large amounts of manganese has been shown to damage liver and kidneys [10-13] as these are involved in its metabolism and excretion. For the same reason liver disease can also lead to impaired clearance of Mn from body increasing the levels of the metal in the body.

Cognitive function is reported to improve after removal of exposure, although motor, sensory and mood disturbances may remain or progress but the timeframe for these events is unclear.[14-17] Thus, the prevention and cure is said to lie in the minimization or cessation of exposure. The US NIOSH recommendation for the occupational exposure limit is 1 mg/m^3 measured as an 8-hour time weighted average and 3 mg/m^3 for short term exposure equal to or less than 15 min. As a safe level for the general public, the US Agency for Toxic Substances and Disease Registry recommends an Inhalation Minimal Risk Level (MRL) of 0.04 $\mu g/m^3$ for chronic exposure.

The US Department of Labour indicated the lowest observed adverse effect limit (LOAEL) to be 0.05 mg/m^3. The same threshold for neurological deficits related to Mn exposure was found by a recent meta-analysis based on individual participant data.[18] The exposure assessment however, is difficult with Mn, because the particles size and shape vary. Thus, the adverse effects may occur with lower levels of exposure in welding because of smaller particle sizes than in mining with dusts with bigger particle sizes. In the literature, however, authors describe exposure in various ways such as total Mn, soluble Mn, inhalable Mn and respirable Mn. These issues have not always been taken into account in previous meta-analyses.

Studies showing severe adverse health effects from very high exposure to manganese have been conducted as early as the 1970s.[19-21] The earliest adverse effects in the nervous system, before overt clinical manganism with Parkinson-like features appears, may appear as poor neurological function as measured by neuropsychological tests such as finger tapping. These effects are often called 'sub-clinical' meaning that they are still at a stage that patients or workers will not notice them. The implicit assumption is that, if exposure continues these symptoms will continue to the stage where one would diagnose overt manganism but this transition from sub-clinical to clinical disease has not been shown. Poor performance on motor function tests may start appearing with exposures at mean concentrations ranging from 0.05 mg/m^3 to 0.30 mg/m^3 of inhalable Mn. However, a relationship between exposure in air and corresponding blood levels or between either of these and adverse health outcomes is still unclear.[18]

Exposure at work exists to variable extent in jobs such as welding and in the steel industry because Mn is an essential part of the process or the equipment used, and thus harmful effects may occur in these occupations to various degrees depending on exposure levels.

Since usually exposure-related effects are dose-dependent, certain low-levels may not be harmful. Therefore, correct knowledge of exposure levels without health effects is essential both to the work places and to the workers in order to protect workers adequately. Many countries enforce low occupational and environmental exposure to manganese by policies stating that the employer should ensure that the work exposure does not cause illness to the employees.

Despite extensive research on Mn exposure, there are still unanswered questions such as what range of health effects may occur, at which exposure levels, and after how long. Furthermore, strength of these relationships are still unclear. Then there are many competitive individual factors that determine the outcome and prognosis of the toxic effects including but not limited to genetic profiles. Finally yet importantly, many times the exposure is not to Mn alone but also to the effects of other exposures in the same industries or non-occupational factors such as liver disease or iron deficiency that are difficult to rule out.

This review builds on the previous work in this area. We followed an a priori protocol so that data driven analysis is avoided and we can answer the questions of interest reliably. The protocol can be seen here: http://osh.cochrane.org

In order to prevent the exposure to manganese and protect workers from resultant occupational disease and disability it is important to find out risk levels and competing factors. When exposure has nevertheless happened, it is also important to have clear criteria for the occupational origin of the symptoms as a basis for financial compensation. A systematic appraisal of the literature should hence result in clear findings about: types or nature of exposure, health effects of exposure, competitive factors and prognosis after stopping of exposure.

The objective of the project is to summarise and assess the evidence on potential adverse health effects of manganese for workers. The goal is to give a descriptive summary of the available evidence about the exposure, its consequences and possible competing factors, as well as to quantify the evidence in a exposure-dose-response relationship and to judge the quality of the evidence.

For this purpose, we took stock of the existing reviews and meta-analysis on the topic of manganese exposure at work and collated the evidence presented in these with data from suitable primary studies that evaluated causal association of manganese exposure to health effects.

Methods

Inclusion criteria:

Study type

1) Systematic reviews

We started with an overview of systematic reviews. The premise being that if already well-synthesized and up to date evidence exists in the form of systematic reviews and meta-analyses, we should avoid duplication of effort. For the systematic reviews and meta-analyses, we were deliberately inclusive so that the entire range of evidence synthesis on the question was appraised. We therefore included any review that:

a) Specified a question consisting of at least one of the following: population (e.g. steel industry workers), exposure (manganese) and outcome (e.g. neurological function deficits). This means that we excluded reviews where the authors stated that the aim was to review the literature on Manganese without further specification.
b) Clearly stated a method and source(s) (electronic databases/non-electronic) for searching studies. We excluded all reviews that did not specify a concrete search strategy.

2) Empirical, follow-up studies, either cohort or case-control.

In addition to the systematic reviews of effects of manganese exposure upon worker health, we collected also cohort studies (prospective and retrospective) of manganese exposed workers. In addition, we included case control studies where patients diagnosed with a disease were recruited and their previous exposure to Mn was explored as a causal factor. However, we included these only if the Mn exposure assessment happened at the start of follow up for example obtaining Mn air levels (ambient or personal) from historical sampling records. We excluded studies that did not specify Mn exposure as such but that relied on a proxy such as mild steel welding. The reason is that it is difficult to estimate what the real exposure would be and that there is likely to be co-exposure with other metals.

We did not exclude litigation studies as long as these fulfilled our inclusion criteria.

We excluded studies that were cross sectional in nature, in which exposure and adverse health effects were measured at the same moment in time. The cross-sectional studies that compare exposed to non-exposed workers are also often called case-control studies. These are obviously different from the case-control studies meant above. There is general consensus that it is not possible to draw causal inferences from these studies as there is no temporality in these studies. The results of causal associations based on these studies are very likely to be biased due to healthy worker effects, participant selection and lack of exposure information in the past.

Participants

Workers of both sexes and any age were included.

General population studies were excluded unless they specified a subgroup of workers.

Exposure

We included studies only if exposure to Mn was objectively measured and reported.

Excluded: Studies that used job titles or activities as proxy for exposure to manganese (JEM, occupational codes, job titles, self-reported exposure history). Even though these types of studies can have a signalling function for possible health risks associated with exposure, we consider the exposure assessment too imprecise to be used for causality.

Comparison

We included studies that compared workers occupationally exposed to manganese to unexposed workers, or those that compared workers exposed to high levels of manganese to workers exposed to low levels.

Outcome

We included studies that measured any adverse health effect.

Often studies presented the results of the neurological tests in more than one format. For these tests, we always chose the averaged measurements for a test over unilateral measurements, and right hand measurements over left hand measurements, when available. Similarly, based on previous research we chose slow finger tapping test (and other similar tests) over fast, when both were presented [22].

Searching and including studies:

We searched multiple databases between 6th and 9th May 2014 and non-electronic sources (reference searching, expert contact) for finding studies without date or language limits. We developed a sensitive search in PubMed and then translated it to the other databases (ToxNet, Inchem, EMBASE, and OSHUpdate) to locate all relevant systematic reviews on the topic to date and all empirical studies. Appendix A. Since many studies only refer to welders and may contain potentially relevant studies on Mn exposure specifically, we developed a supplementary search for welders and the results of this were added to those from other searches.

Two reviewers (SI, JV) independently checked fulfilment of the inclusion criteria first via titles and abstracts and then via full text. Disagreements were resolved by discussion.

Data extraction and management:

Two reviewers (SI, JV) independently extracted data from the studies that were included. The detailed forms used for both the overview and systematic review are provided in Appendix B.

Risk of bias assessments

We assessed the risk of bias in the included primary studies by adapting a checklist for assessing the quality of observational studies as proposed by Shamliyan.[23 24] According to their proposal, we assessed six major domains of internal validity: exposure assessment; outcome assessment; masking of assessors; confounding factor (competitive factor) adjustment; attrition; and analysis methods. We supplemented this with three minor domains of bias, which we found relevant for this particular question: ethical approval of the study, funding for the study, and conflict of interest in the study. We considered conflict of interest an important item because both workers and industry have financial interests in the results of research into the effects of manganese.

For domains with more than one subcategory (for example outcome assessment involves the source of data, the definition of outcome and the measurement of data) a low risk judgement was given to the domain if all subcategories were scored low risk. A moderate risk judgement was given to a domain when all subcategories were marked moderate risk or no more than one of these was marked high risk. An overall high risk for a domain was given when more than one subcategory were marked high risk.

We used the same principle for giving a study an overall risk of bias: low risk=all domains at low risk; moderate risk=all domains at moderate risk or only one domain at high risk; and high risk= more than one domain at high risk.

For included systematic reviews, we used the AMSTAR tool to assess risk of bias.[25-27] See appendix B for the full checklist. An ideal review was to score 11 on the AMSTAR, however, this was deemed unlikely considering the state of evidence synthesis outside intervention reviews in general and that AMSTAR is relatively recent and has not yet been adopted universally. We decided that a score of 5 or above, depending on existence of major flaws, would be enough to categorise a review as good quality and therefore we would restrict our implications to these reviews alone.

All assessments were done independently by two authors (SI, JV) and consensus reached by discussion.

Data Analysis:

First, we analysed the systematic reviews on exposure assessment to better be able to define the exposure. Then, we analysed the adverse health effects that were evaluated in the studies found and listed those that were most probably related to Mn-exposure. Based on these findings, we conducted an overview of the included systematic reviews, complemented with evidence synthesis from primary studies that fit the inclusion criteria.

Adjustment for confounding

We predefined important confounders (competing factors) that should be adjusted for when studying the effect of Mn exposure on health. These were: age; sex; socioeconomic status; education; alcohol; smoking; iron status and liver health. Increasing age is associated with poorer neurological function. Females perform better on some neurological tests and worse on others compared to their same age male counterparts. Smoking has a protective effect against some neurological adverse effects. Poor iron metabolism and liver function increase the accumulation and thus toxicity of manganese. Education and socioeconomic status not only influence what job one works but also the cognitive function tests with higher education and status associated with better neurological function. We had planned to adjust for the important unadjusted confounding factors within each study following the methods described by Greenland.[28 29] This was not possible because suitable data were not available on these competing risk factors for any of the included studies.

Dealing with missing data

We had planned to impute standard deviations when these could not be obtained from reports or unpublished data from authors. This was possible for only one study (Blond) where standard deviation for the mean difference between exposed and control groups could be calculated from the reported *p* values.

Previous clinical research indicates no clear pattern between increasing exposure and the development of symptoms due to manganese. Therefore, we had planned to assume and test a linear relationship between the natural logarithm of RR and increasing exposure. For this purpose, we considered the exposure to 0.05 mg/m3 as a threshold for this review and 0.01 as the incremental step of increased exposure for inhalation based on literature. [18]However, due to lack of data this was not possible.

Data synthesis

We planned to add the results from any additional primary studies to the findings of the overview. However, this was not possible due to the heterogeneous methods of the included reviews. Thus, we report findings from the primary studies separately by conducting a new evidence synthesis and meta-analysis where possible. Alternatively, the findings were tabulated with narrative analysis. In meta-analysis, we presented risk ratios or mean differences as estimates of the effect. Because adverse health outcomes were infrequent, we used ORs equivalent to RRs. We analysed each disease outcome separately, or if suitable, in subgroups.

We could not use RR per unit increase in exposure, as these data (effect for exposure in increasing doses) were not available for pooling. We combined the RRs for total exposure for a disease outcome using the most adjusted natural logarithms of the relative risk as input for a random effects meta-analysis. We analysed mean and standard deviations for continuous outcomes of neurological function, such as test score for finger tapping, presenting our results as mean difference and standard error.

At first, studies we assessed studies for similarity of participants, exposure (route, duration and intensity, follow up) and outcome measurement and grouped for analysis accordingly. We wanted to sub-group exposures by job/occupation/industry. However, this was not always possible as we found very few studies.

Next, we assessed statistical heterogeneity in the meta-analyses by means of the I^2 statistic. We considered an I^2 value of up to 25% as low, values between 25% and 75% as moderate, and values over 75% as high degrees of heterogeneity respectively.

Assessment of reporting biases

We avoided language and publication bias by including studies in any language and of any publication status. Later we had planned to assess publication bias by observing a funnel plot and applying Egger's test to the included studies, however, too few studies were found for either of these tests to be meaningful therefore these were not done.

Subgroup analysis and investigation of heterogeneity

We planned to evaluate if the outcomes varied according to:

- the types of occupation, to analyse workers by industry.
- the year of the study, to separate studies carried out before and after the year 2000.
- the country of the study, to differentiate between study participants from Western Europe and the US versus study participants from Asia.

However, these could not be done due to few studies.

Sensitivity analysis

We had planned to evaluate if our results were sensitive to the inclusion of low quality studies with a high risk of bias, by excluding high risk studies from the meta-analysis. However, very few studies were located and no study scored a low risk of bias. No sensitivity analyses planned could be performed due to few studies.

Grading and Strength of causality of the evidence

We used the approach of the Danish Occupational Medicine Association to grade the strength of causality into one of the following five categories:

1) +++ strong evidence of a causal association
2) ++ moderate evidence of a causal association
3) + limited evidence of a causal association
4) 0 insufficient evidence of a causal association

5) evidence suggesting lack of a causal association

Description of categories:

Strong evidence of a causal association (+++):

A causal relationship is very likely. A positive relationship between exposure to the risk factor and the outcome has been observed in several epidemiological studies. It can be ruled out with reasonable confidence that this relationship is explained by chance, bias or confounding.

Moderate evidence of a causal association (++):

A causal relationship is likely. A positive relationship between exposure to the risk factor and the outcome has been observed in several epidemiological studies. It cannot be ruled out with reasonable confidence that this relationship can be explained by chance, bias or confounding, although this is not a very likely explanation.

Limited evidence of a causal association (+):

A causal relationship is possible. A positive relationship between exposure to the risk factor and the outcome has been observed in several epidemiological studies. It is not unlikely that this relationship can be explained by chance, bias or confounding.

Insufficient evidence of a causal association (0):

The available studies are of insufficient quality, consistency, or statistical power to permit a conclusion regarding the presence or absence of a causal association.

Evidence suggesting lack of a causal association (-):

Several studies of sufficient quality, consistency and statistical power indicate that the specific risk factor is not causally related to the specific outcome.

The classification does not include a category for which a causal relation is considered as established beyond any doubt. The key criterion is the epidemiological evidence. The likelihood that chance, bias and confounding may explain observed associations are criteria that encompass criteria such as consistency, number of ‘high quality’ studies, types of design etc. Biological plausibility and contributory information may add to the evidence of a causal association.

In addition, we used the GRADE approach to assess the overall quality of evidence. A formal assessment of the quality of evidence in included reviews was done by two reviewers according to the widely recognized GRADE approach. The GRADE approach specifies four levels of quality:

- High quality for double-upgraded observational studies.
- Moderate quality for upgraded observational studies.
- Low quality for double-downgraded observational studies and
- Very low quality for triple-downgraded observational studies; or case series/case reports.

Randomized evidence is unfeasible in work place exposures assessment, not to mention unethical, unless a preventive intervention was being tested for exposure reduction. We considered the best evidence to be from prospective follow up studies that controlled well for confounders of age, exposure dose, smoking habit, education level and alcohol habits. Using GRADE then, we downgraded evidence by one or two levels from this well-controlled prospective cohort study depending on the presence of five factors:

Serious (-1) or very serious (-2) limitation to study quality

Important inconsistency (-1)

Some (-1) or major (-2) uncertainty about directness of evidence for causality

Imprecise or sparse data (-1)

High probability of reporting bias (-1).

Similarly, we upgraded evidence based on lack of confounding (+1), a large effect size (+1) or a clear exposure-dose response (+1) seen.

Results

Search results

Our search yielded 9752 references and after duplicate assessment identified 42 relevant citations of potential systematic reviews or meta-analyses and 91 citations for primary studies on the basis of title or abstract. These were then reviewed in full text along with their reference lists. This led to inclusion of a set of 15 papers[18 30-43] referring to 13 systematic reviews/meta-analyses, and 13 papers[14 16 44-54] referring to eight primary studies as per our inclusion criteria. The excluded reviews and primary studies are listed in tables in appendix C with reasons. Two ongoing studies were identified.[32 55].

Among the primary studies 30 were not found in full text till the submission of this report. These are listed in appendix C also.

The figure 1 below shows our search and inclusion assessment process in detail.

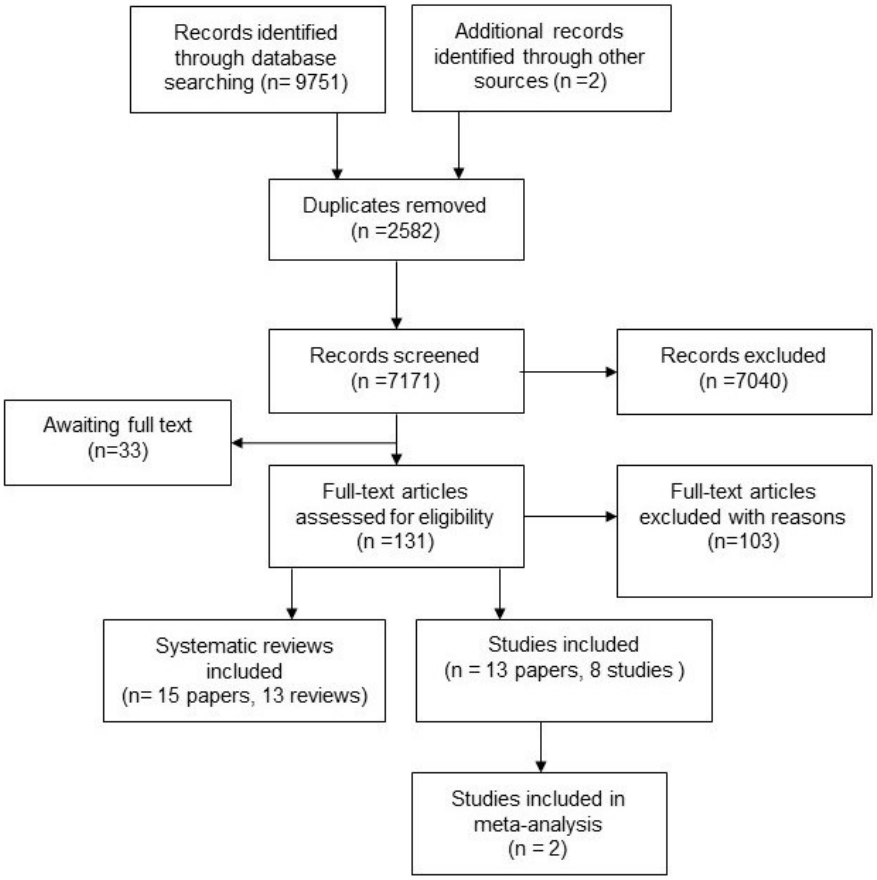

Figure 1PRISMA flow chart of study inclusion

Characteristics of included reviews

The included reviews were spread over the past decade with the first one published in 2004 and the latest in 2014. Nearly half (n=6) were published between 2009 and 2011.

Nine reviews [18 31 33-37 40 43] dealt exclusively with working populations, two of these[34 40] limiting the working population to welders. The remaining four [30 32 41 42] included populations in addition to workers for example communities, children, general adult population or even animals and bacterial cells. These reviews provided results for occupationally exposed samples separately.

The exposure most often assessed was Mn in air (n=11) followed by Mn in the human tissues (blood, urine or brain) (n=6) and as welding fumes (n=2). Exposure ranges evaluated were rarely reported, and where reported these varied widely with lowest reported exposure at 1.0 $\mu g/m^3$ [32] to the highest at 10.58 mg/m^3 [37]. Durations were reported even less often (n=2), and ranged between 5 and 12 years, [31] and 5 and 21 years. [18]

The included reviews often did not specify to which type or intensity of exposure the exposed group was compared. However, based on information within tables of included studies the most often addressed comparison was Mn exposure versus an assumed non-exposure (n=5). Two reviews compared varying levels of Mn exposure to assess a exposure-dose related effect. [32 34]

The reviews addressed various aims and therefore a range of outcomes. These included exposure outcomes as well as health outcomes and included the following: manganese biomarkers (n=3); optimal test of neurological effects of manganese exposure (n=2); the confounding effect of demographics on health effects of manganese (n=1); and the adverse health effects of Mn exposure (n=7).

Four reviews [36 37 40 41] restricted their inclusion to follow up studies, while the rest [18 30-35 42 43] included, either exclusively or largely, cross sectional studies.

The detailed characteristic of the included reviews is provided in appendix D

Characteristics of included primary studies

Eight empirical studies were included, all published between 1999 and 2014.[14 16 44 46 49 50 52 54] Four were from Northern Europe and one each from the USA, Canada, South-Korea, and Iran. Four were prospective and four retrospective follow up design (see table 1). One study did not report a follow up duration, however, the follow up ranged from 1 to 29 years for the rest. The study size varied: two studies were small with less than 30 total participants; three were medium sized with samples sizes of 74, 135 and 201; three studies included several thousand participants each.

Participants were miners in one study, welders in three studies, while the rest included factory workers (steel industry / battery industry). All studies were exclusively done in men and the mean age, when reported, ranged from 26 to 56 years.

Exposure to Mn was measured for the purpose of the research in only two studies [44 54] whereas in the rest of the six studies the data on exposure levels was obtained from existing administrative databases. Exposed population was defined in three studies as a minimum duration of exposure at work [14 50] and the inclusion threshold ranged from six months to 1 year, while two studies did not report any threshold with the exposure. [44 46]Three studies did not report any definition.[16 49 54] Mean exposure levels for Mn ranged from very low (0.008 mg/m^3) to very high (64mg/m^3), but for all studies except the Iranian study on miners,[46] the exposure was below 1mg/m^3. This means the results of these studies apply to exposure below this level. Only three studies provided values of a mean composite exposure index (CEI) in mg/m^3years. [14 16 46]

Data driven categorisation of exposure from low to high intensity was provided in three studies.[16 52 54] Fored[49] categorized Mn exposure by types of welding. Blond[44] separated exposure levels into two categories: before the 90s; and after the 90s. Bowler separated exposure levels of air Mn from blood Mn levels and CEI levels. Hobbesland[50] categorized exposure as that in furnace workers and in non-furnace workers

Reference group was not defined in any of the studies, but it was clear from two studies that the reference group consisted of workers from the same region but from outside the exposed site.

The most common outcome in the studies was a variety of neurological function tests (n=4).[14 16 44 54] Two studies evaluated the incidence of ICD classified movement disorders such as Parkinson's' disease.[49 52] One study evaluated lung function and the incidence of occupational respiratory diseases.[46]

Studies often assessed the confounding effect of age (n=6) and education (n=5) on the effect of exposure, followed by alcohol (n=4) but only one [54] assessed the effect of any medical pathology.

Table 1 Characteristics of included primary studies of health effects of Mn exposure at work

Study ID	Blond 2007	Boojar 2002	Bouchard 2007	Bowler 2011	Fored 2006	Hobbesland 1999	Park 2006	Roels 1999
Country	Denmark	Iran	Canada	USA	Sweden	Norway	South Korea	Belgium
Study design	Prospective cohort	Prospective cohort	Retrospective cohort	Prospective cohort	Retrospective cohort	Retrospective cohort	Retrospective cohort	Prospective cohort
Duration of follow up years	1-28 years	7 years likely	14.4 years	15 (10.9)	29 years	3 - 20 years	NR	10 years
Exposed participant source	Steel plants	Iron Mines	Ferro alloy plants	Bridge construction site	Swedish national census database	Ferro alloy plants	Ship building company	Dry cell battery plant
Participants analysed 'n'	Exp=60; Ref=14; Total=74	Exp=145; Ref=65; Total=210	Exp=68; Ref=67; Total= 135	26	Exp=49488; Ref=489572; Total=539,060	Total=6,363	Exp= 24,963; Ref= 13,597; Total= 38,560	24
Mean age years	Exp=56; ref 54	26 (5.3)	Exp=44.3; Ref=43	43(9.2)	NR	NR	NR	31 (22-50)
Gender	Men	Men	Men	Men	Men	Men	Men	Men
Attrition	Exp=46%; Ref=63%	NR	Exp=33%; Ref=44%	46 %	NR	NR	NR	60 %
Occupation/ industry	Plant workers/steel	Miners/ iron mining	Plant worker/steel	Welder/construction	Welder/any	Furnace workers/steel	Welder/ shipbuilding	Factory workers/ battery industry
Exposure data source	Industrial hygiene data from 1990s	Personal air sampling done for the study	Sampling data from company records+ interviews	CalOSHA air exposure data for the site	Industrial hygiene survey 1974/75	Personal samples likely from company records	Company records of yearly exposure levels plus job duration	Personal air sampling done for the study
Exposure categories	1970-1990;>1990	NR	Lowest tertile; Middle tertile; Highest tertile; all	MnA; MnB; CEI	Mild steel MMA; Mild steel MAG; railroad MMA; Stainless steel MMA	Furnace workers, non-furnace workers	High; Low; Very low	Low; Medium; High
Exposure definition	Mn in air; no threshold reported	Miners; No threshold reported	NR	Welders who worked 1 to 2 years at the site	NR	Six months employment at any of the four plants between 1933 to 1991	At least one year employment as welder at companies between 1970 and 2000	NR
Air exposure(SD)mean/median/ range mg/m3	1990= 0.03	64 (41)	NR	0.2 (0.08)	Mild steel MMA=0.26; Mild steel MAG =0.3; railroad MMA=0.13; Stainless steel MMA=0.14	0.99	High=0.88(3.1); Low=0.1 (4); Very low 0.008	1985 value TWA for Mn total dust: 0.948 1985 value for subcategories: Low= 1;Medium=3.3; High=9.7 1995 value for subcategories: Low= 0.5 Medium: 0.7 High: 3
CEI mg/m3 Year	NR	At start=1365(645); at end= 114(66)	27.38(24.7)/ median= 19.03; range 0.3-100.2	0.21	NR	NR	NR	NR
Exposure in the reference/control group	NR; who did not work in steel plants (office, smiths, electricians)	NR; not miner workers	NR	NR	NR	NR	NR	NR

Confounders controlled for	Alcohol; Lead Exposure	Age; Salary; Education; Smoking;	Age; Education; Alcohol	Age; Education; Ethnicity; Total Years As Welder	Age; County Of Residence; Education; Time Period	Age, Employment Duration	Age	Age, Smoking, Alcohol, Socioeconomic Status, Education, Medical Pathologies, Hyperthyroidism, Depression, Coffee
Outcomes: measure	Neurological tests (finger tapping; reaction time) CATSYS scores	Incident occupational asthma/ bronchitis= self-report+ clinical diagnosis from medical records; FEV FVC test scores	Neurological tests with computer and questionnaire/interview	Neurological tests with CATSYS and clinical assessment	ICD codes G20 to G26 (PD/basal ganglia disease)	Cancer incidence as per national database	ICD codes G20 to G26 (PD/basal ganglia disease)	Neurological tests for eye hand co-ordination , hand steadiness and tremor intensity, and simple visual reaction time in 10-2 sec/ 30 sec.

Legend: CATSYS= A Computerized Test Battery http://www.catsys.dk/; CEI= Cumulative Exposure Index; Exp= Exposure group value; FEV= Forced Expiratory Volume; FVC= Forced Vital Capacity; ICD= International Classification of Diseases; MMA= Manual Metal Arc; MAG= Metal Active Gas; MnA= Manganese in Air; MnB= Manganese in Blood; NR= Not Reported; PD= Parkinson's disease; Ref= Reference group value; TWA= Time Weighted Average;

Risk of Bias

Risk of bias in included reviews

Only two reviews scored above five on AMSTAR, our a-prior but arbitrary threshold for the reviews' quality to be satisfactory.[36 41] There was no time trend showing better quality indicating increased uptake of guidelines in the more recent years. Although AMSTAR rates all questions equally, we believe the more important predictors of review quality are presence of a protocol and duplicate study selection and data extraction. This was also evident from our results where the only two reviews scoring higher than five were also the only ones performing these tasks in duplicate. The search was comprehensive in seven of the 13 reviews which increases confidence in not having missed any important piece of evidence. However, a flow diagram of how studies were included was provided by two reviews only,[36 41] and a list of excluded studies was provided by only two reviews also.[18 41] A list of excluded studies helps a decision maker in understanding why, which and how many studies were not included. In our experience the list of excluded studies in epidemiological reviews is often bigger than the included study list, and can be provided as an online appendix if space is a concern. Only one review assessed and incorporated individual studies' quality into their analysis and only one assessed publication bias. None of the reviews reported conflict of interest sufficiently.

Nature of exposure

Six reviews provided information on the nature of Mn exposure in air, its range and its relationship with the welding fumes or biomarkers.[18 31 32 34 36 43] Only one of these scored over five on AMSTAR [36] providing good quality synthesis of evidence. Risk of bias in included studies was not assessed in any review.

The Hobson [34] review acknowledged that some data were missing and also that varied or undefined methods of measurement were used in included studies, but these were not accounted for in their analysis or conclusion. The lack of any quality assessment of the included evidence also diminishes the reliability of these model based findings for an accurate representation of prevalent exposure levels. With a search limited to one database, no quality assessment of included studies, unclear criteria for suitable data for the calculation of NOAEL, and no list of excluded studies or reasoning for it, we considered that the narratively synthesized findings from Bailey [31]are at a high risk of bias.

The exposure and outcome data measurement in studies were not appraised for reliability in the IEH report[43], nor any other quality indicator, even though general comments on poor quality of data were made. Finally yet importantly, the exposure assessment was different in the three studies included by this review (respirable versus total Mn) and the design of all three studies was cross-sectional, with one using no control group. This makes the reported NOAEL values less reliable. It could well be that the actual value is much below the one recommended.

For the relationship between manganese in air and that in the body (biomarkers of exposure), although no formal quality assessment was done in the Baker review[32], the authors did identify the limitations of the existing studies and provided the plan and preliminary results of a large well planned study to address these limitations. We believe it

would be best to wait for the results of this new primary study to come out, as these may provide the most valid answer for this question.

In terms of methods, Li et al.[36] used more stringent criteria for study designs included, reported a more comprehensive search along with duplicate study selection and data extraction as per the accepted guidelines. This resulted in better quality rating as well as clear conclusions. However, they did not incorporate the quality assessment they performed for included studies into their findings. It should also be noted that the review correlated Pallidal index to Mn in blood, not with Mn in air which is the actual exposure variable.

Adverse health effects

Seven reviews addressed one or more health outcome for association with Mn exposure.[18 30 33 35 37 40 41 43] The only good quality review (AMSTAR 7) was on the sporadic amyotropic lateral sclerosis.[41] This was also the only review that assessed the quality of the studies and incorporated this into their conclusions. Thus their findings are considered valid and reliable. Of the rest of the six reviews evaluating a casual effect of Mn exposure at work on health outcomes, we consider the two scoring 4 on AMSTAR [40 43] as moderate quality. The remaining five did not reach that standard.

Overall risk of bias

Thus the overall level of the existing systematic reviews and meta-analyses is low with a high risk of bias of their findings. Four reviews or their primary authors were sponsored by the industry. This further raises the likelihood of bias in the selection and presentation of data. The table 2 below provides the AMSTAR assessment summary of the included reviews.

Table 2 AMSTAR scores in included reviews

Review ID	Assem, 2011	Bailey, 2009	Baker, 2014	Greiffenstein, 2007	Hobson, 2011	Lees-Haley, 2006	Li, 2014	Ma, 2011	Meyer-Baron, 2009	Mortimer, 2012	Sutedja 2009	Zoni, 2007	IEH, 2004
1. Was an 'a priori' design provided?	0	0	0	0	0	0	0	0	0	0	0	0	0
2. Was there duplicate study selection and data extraction?	0	0	0	0	0	0	1	0	0	0	1	0	0
3. Was a comprehensive literature search performed?	0	0	0	1	0	1	1	1	0	1	1	0	1
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?	0	0	0	0	0	0	1	0	1	1	0	0	1
5. Was a list of studies (included and excluded) provided?	0	0	0	0	0	0	0	0	1	0	1	0	0
6. Were the characteristics of the included studies provided?	1	1	1	1	1	1	1	0	1	1	1	0	1
7. Was the scientific quality of the included studies assessed and documented?	0	0	0	0	0	0	0	0	0	0	1	0	0
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	0	0	0	0	0	0	0	0	0	0	1	0	0
9. Were the methods used to combine the findings of studies appropriate?	0	1	1	1	1	1	1	1	1	1	1	1	1
10. Was the likelihood of publication bias assessed?	0	0	0	0	0	0	1	0	0	0	0	0	0
11. Was the conflict of interest included?	0	0	0	0	0	0	0	0	0	0	0	0	0
Total per review	1	2	2	3	2	2	6	1	3	4	7	1	4

Risk of bias in primary studies

Of the major domains of internal validity, the assessment of outcome was the domain with the lowest risk of bias over all, with five studies at low risk and the remaining three at moderate risk of bias.

Unlike outcome assessment, exposure assessment was marked consistently at moderate risk across studies. The most common finding across the studies was a lack of definition of exposure or non-exposure. Thus all the studies divided workers into exposed and non-exposed without any actual measurements of Mn being performed on the referents while only two [16 50] defined what they considered exposed. When they did, the definition was always about the length of exposure as measured by number of years on the job.[16 50] Even when cumulative indices were calculated, these were not used to either categorise exposure into research informed categories of high or low exposure but were often not analysed for the outcome. Almost all studies reported a mean ambient or personal air level and these varied from 0.03 to 64 mg/m^3 of Mn but without including this in the definition of exposure. Another important fact was that none of the studies measured Mn levels in the control or referent groups. It has been seen that near the exposure sites such as factories handling manganese the ambient air exposure to Mn for general population may also be higher than the recommended values. (Moreno, 2011 #10490}

Masking of assessors was reported in one study only. This is an important piece of information when analysing cause and effect relationships because the information can bias the assessor either way although not necessarily in a random fashion. For example, the effect of sponsors and an assessor's prior involvement with either side of the argument might bias their categorization. This may not be true for certain outcomes such as cancers where a histopathology report decides the incidence. However, for the subjective outcomes of neurological performances this does matter, especially when authors exclude data of participants because the authors believed these participants were malingering [56]. Similarly, prior knowledge of outcome in a retrospective cohort analysis can bias the way exposures are assigned. The absence of protocols for risk factor studies further limits our ability to judge the bias present.

The complete set of seven confounders that we considered important for the relationship between Mn and adverse health outcomes was: age; education; socioeconomic status; liver and iron status; smoking, and alcohol. Gender was not considered important since the studies are almost exclusively in males, and the few females are usually excluded from analyses. Another important factor in our opinion would be the compliance with using safety equipment; however, no data on this factor were available to make a judgement. Only one of the included studies assessed all the seven important factors [54] and was at a low risk of bias in this domain. The other most adjusted studies were Bouchard [16] and Boojar [46] each adjusting for four of the total set each.

Attrition was the most at risk domain with half of the studies not providing information and the other half at high risk with nearly half of the sample lost to follow up. This is an issue requiring care in the future studies where a priori anticipating the loss to follow up and planning for this may help.

Methods of analysis were found largely to be at moderate risk of bias with only one study where we could not make the judgement: Boojar[46] did not provide an analysis of their data but simply reported the initial and final values per group. Therefore, although it still allowed us to do an analysis ourselves, we consider it poor reporting and therefore marked it unclear for this domain. For the other studies, it was surprising to see no study at low risk. For a good quality epidemiological cohort study, we expected a sample determined and justified a priori for a certain hypothesis tested. Then, ideally, we would have expected to see methods planned at the start to counter baseline imbalances often associated with non-randomized designs. An example is matching for confounders at sampling stage and then adjusting/ testing these in analysis. We also expected a fully planned exposure-dose-response analysis from studies if they planned to find a harmful level of exposure causing an effect of Mn, because an arbitrary cut off or studying just one level as a mean would not identify what level becomes harmful and in which individuals. Finally, we expect all aspects of an analysis presented - from means and standard deviations to numbers analysed- while accounting for attrition. These criteria were not met by any study.

Of the three minor domains funding was low risk for most studies as these were either completely grant funded or co-funded by grant(s) and the industry, with one study also reporting that the funding organizations had no input in the study from start to finish.

Conflicts of interest were reported in two studies and were not reported in others. This may reflect better reporting standards in some journals. For all except one study[46], we deem the risk of an ethical conflict to be low.

Overall, for internal validity domains, no study had more than one domain at low risk of bias and therefore we consider none of the studies to be at low risk of bias. All of these are at moderate risk, with no more than one domain at high risk of bias. Table 3.

Table 3 Risk of bias in the included primary studies of the health effects of Mn exposure at work

Study ID	Blond 2007**[45]**	Boojar 2002**[46]**	Bouchard 2007**[16]**	Bowler 2011**[14]**	Fored 2006**[49]**	Hobbesland 1999**[50]**	Park 2006**[52]**	Roels 1999**[54]**
Ethical approval	Low risk	High risk	Low risk	Low risk	Low risk	Unclear risk	Unclear risk	Unclear risk
Funding	Low risk	Moderate risk	Low risk	Moderate risk	Low risk	Low risk	Low risk	Moderate risk
Conflict of interest	Unclear	Unclear risk	Low risk	Unclear risk	Low risk	Unclear risk	Unclear risk	Unclear risk
Outcome assessment (case definition and measuring)	Low risk	Moderate risk	Low risk	Low risk	Moderate risk	Low risk	Low risk	Moderate risk
Exposure assessment (case definition and measuing)	Moderate risk	Moderate risk	Moderate risk	Moderate risk	Moderate risk	Moderate risk	Moderate risk	Moderate risk
Masking of assessors (exposure and outcome)	Unclear	Low risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk
Confounding factor adjustment/ basline imbalances	Moderate risk	Moderate risk	Moderate risk	Moderate risk	Moderate risk	High risk	Unclear risk	Low risk
Attrition	High risk	Unclear risk	High risk	High risk	Unclear risk	Unclear risk	Unclear risk	High risk
Analysis (methods, justification, reporting of results)	Moderate risk	Unclear risk	Moderate risk	Moderate risk	Moderate risk	Moderate risk	Moderate risk	Moderate risk

Review findings

We present our results for the range and nature of Mn exposure first, followed by the nature of outcomes assessed in literature to date. For these two concepts, systematic reviews and meta-analyses of Mn exposure were considered, as planned. Then we present a synthesis of evidence on the main question of causal relationship of manganese exposure to adverse health effects. For this, we complemented the findings of the systematic reviews and meta-analyses with our own search for primary data on the causal link.

Exposure

For this component our main questions were:

- What is the nature and range of Mn exposure in workers?
- Is there a relationship between manganese air exposure levels and biomarkers in blood, urine or brain?

We found three reviews addressing the first question [31 34 43]and four addressing the second [18 32 36 43]. The characteristics of these are presented in appendix D.

Nature and range of exposure

1) Relationship of welding to manganese exposure

Only one review Hobson 2011 [34] (Appendix D) assessed the relationship between welding and consequent manganese exposure levels. The review included only English language published literature and modelled all welding particulate exposures reported as independent variables with Mn levels as dependent variable. The review concludes that manganese compounds make up about 4% (Mn = 0.03–1.19 mg/m^3) of welding fumes particulate mass. The model derived from the publications could explain around 75% of the variance in exposure by using the type of welding and if welding was done in a confined space or not. No time trend in exposure reduction was seen, however, the independent variable (welding fumes) was largely based on publications prior to 1999 and the dependent variable (Mn) on publications largely after the year 2000 and therefore may not correspond exactly as exposure limits were lowered in the mid to late 90s in many parts of the world.

2) Safe exposure levels of manganese at work

Bailey [31] and IEH 2004 [43]were the two reviews that attempted to calculate a safe exposure level for Mn in the form of an NOAEL value.

Bailey [31] used comparative studies assessing neurological function outcomes to observe no effect levels. Based on three of the 12 identified studies they concluded that the human NOAEL is 60µg/m^3 respirable Mn, an average of the exposures in the three studies.

The other review that provided an NOAEL was the IEH report [43]. The review searched multiple (n=16) databases and tabulated 13 studies for the question of prevalent exposures in the work places. They noted that Mn levels varied from site to site, within a site at various times and between inhaled fraction and total ambient dust. Furthermore the ambient values may differ from personal air samples. The report preferred the arithmetic mean over geometric mean of exposure levels and suggested an NOAEL of 1 mg/m3. The report also advised limiting exposure to 0.1 mg/m3 respirable manganese. Their tabulation showed a range from 0.03 to 114 mg Mn/m3 for the 13 studies including both arithmetic and geometric means. In the presence of such variation it is not possible to be confident about the values of NOAEL suggested by the report (NOAEL 1 mg/m3; limiting exposure to 0.1 mg/m3 respirable manganese). The author used the three 'negative' studies, which found no adverse health effects, as the basis for this NOAEL value, however, it is not clear how they reached this figure. The average of the exposure values in these studies is 0.42 mg/m^3. It could be that they chose the lowest exposure value of the three studies which is 0.1 mg/m^3. What however remains unexplained is the much lower exposure values in some other studies in the same table which apparently were not 'negative studies'.

Relationship between manganese in air to manganese biomarkers

Of the four reviews addressing this question, one [18] did not directly assess the relationship between Mn in air to Mn in blood, but indirectly compared effect slopes obtained for Mn in air versus results for all neurological tests and for Mn in blood for the test that were positively related to Mn in air (digit symbol, finger tapping). Based on these comparisons the authors concluded that Mn in blood was not well correlated with Mn in air.

The IEH report[43] narratively reported results of 27 reports of either blood (Mn-B) or urine levels of Mn (Mn-U) as biomarkers of exposure. The conclusion was that the variation in Mn-B and Mn-U even in the unexposed worker is too much to reflect exposure accurately. Also, the correlation between air exposure levels and the blood or urine levels, in all studies that presented this, was poor.

The review directly correlating Mn in air to Mn in blood was Baker [32]. The review found that Mn in air and blood correlated only after the exposure in air increased beyond 0.01 mg/m^3. However, a clear relationship between the two could not be found. The review also separately analysed welder studies but for these the air exposure levels could not be established due to lack of data.

The review by Li et al. [36] correlated Mn in blood to a biomarker for brain accumulation of Mn – the Pallidal Index (PI). Li et al. who limited their data for exposure levels above Mn concentrations of a time-weighted average of 0.15 mg/m^3 found that PI corresponds well to the Mn in air.

Adverse Health Outcomes

Given how Mn enters the body and what is known about its physiology from animal and laboratory research, we limited our outcomes to three main organ systems that have been studied as potential sites of toxicity in man. These are the neurological system, the respiratory system, the male reproductive system, and cancer.

Neurological disorders

Manganism

The most well-known effect is the neurological condition called 'manganism'. Manganism is diagnosable as an entity separate from Parkinson's disease. It has the following cardinal features: occupational exposure to manganese or its compounds, generalized bradykinesia and rigidity, intention tremors, dystonia, and cock-gait or slapping broad-based gait. There is, however, no generally accepted definition for clinical diagnosis, though experts have published useful criteria to achieve uniformity.[57 58] We do not know of epidemiological research that has applied these full criteria formally in a study. Confidence in clinical diagnosis increases with MRI and PET scans and tests showing no consistent response to dopaminomimetric drugs and a positive response to ethylene diamine tetrahydrochoride (EDTA) chelation.[59] MRI findings include bilateral increase in signalling confined to globus pallidus and substantia nigra. This increased signalling may dissipate after removal of exposure. Thus, MRI findings may be of little use for previously exposed workers. However, the occupational exposure levels have gone down over the past several decades leading to no cases of manganism in the recent past.[60]

Parkinsonism, or Parkinson's disease

Manganese exposure has been linked to an increase in Parkinson's disease. We included this outcome as reported by authors if a physician made the reported or indicated diagnosis using valid scales such as The Unified Parkinson's Disease Rating Scale (UPDRS).

Neurological tests

From within the range of neurological (behavioural or psychological) test scores that are abundantly used in this area of research, we aimed to restrict our analyses to the most suitable and reliable data. We looked for a) those tests that are congruent with early manganese signs and symptoms; and b) important competitive variables that must be considered when synthesizing evidence on causal association of manganese and neurological effects.
For this purpose, we evaluated the included reviews that addressed these questions.

a. <u>Optimal tests for neurological outcomes of Mn exposure</u>

Two reviews addressed this important question. [18 42]
Zoni used methods that were not entirely clear or appropriate as to how the conclusions were reached (was it that a test found positively associated in 6 or more studies was suitable, or that the top five positively associated test would be

used). The review found the most commonly used and positively associated tests of neurological functions were simple reaction time (11 studies), finger tapping (12 studies), digit symbol (13 studies), digit span (13 studies), and other symptom questionnaires (9 tests). However, in the conclusions the review listed other tests suitable for assessing neurological function (Standard progressive matrices or WAIS-R; POMS or mood scale BSI; visual reaction time, pursuit aiming, pegboard Test; Trail making test, addition test,Rey-15 item, or Benton visual retention or WMS; Tremor test, Luria-Nebraska motor battery; CATSYS system) which was inconsistent with their results.

The review did not differentiate between the various indices of exposure, MnA, MnB, MnU, or CEI. This can be problematic since not all of these markers correlate equally with effects, or with each other. No consideration was given to study quality, size, or confounding in formulating conclusions. Also, since the outcome is the same (biomarkers) whether the exposure is occupational or not, the test would apply the same for both. Thus the reason for separately analysing adult environmental studies was not clear.

The other review[18] assigned the tests to cognitive or motor function domains before analysing these and provided justifications for their choices. For example, that the simple reaction time (SRT) is a more relevant test of motor speed than it is of cognitive function, while its standard deviation may be a better test for the speed of information processing. We think this model can be used in future studies for the same domain/outcome, therefore their list of appropriate tests for each domain were used as primary outcome measures in this review. Based on the model from Meyer-Baron review we considered five domains of neurological function important to be evaluated. These are presented below along with their respective tests (table 4). Although finger tapping is as non-specific as others for neurological damage in general, it can be more sensitive and specific for basal ganglia neurological damage with early signs of parkinsonism [61]. Thus the findings of this test were relied upon more in our conclusions.

Table 4 Domains of brain functions relevant to manganese exposure and their related neurological tests

Domain/ outcome	**Test**	**Description**
Recognition (visual perception and memory)	Benton visual retention	Test consist of 10 cards containing 1 to 3 figures with increasing difficulty. After an exposure of 10 seconds, the task is to reproduce each card by drawing. The number of correct reproductions and errors is recorded.
Short term and working memory	Digit span	Test consists of repeating orally presented digits either forwards or backwards. Digits forward measures attention span, while digits backward involves both a memory component and a reversing operation (mental double-tracking). The test score is the total number of digits correctly repeated forward and backward.
Motor function/Fine motor performance/Dexterity	Finger tapping	The subject is required to press a button with the index finger as many times as possible within 30 s, first with the preferred hand, then the non-preferred hand, and finally with both hands alternately tapping two buttons. The number of buttons pushed in each trial and the number of errors in alternating tapping are recorded.
	Simple reaction time	The task is to press a button as fast as possible when a large square appears on the computer screen. Individual reaction-time latencies are recorded. This test measures visuomotor speed and attention.
	Luria-Nebraska finger-thumb sequential touch	This test is part of the Luria-Nebraska test battery and measures motor function. The patient is requested to touch each finger in turn with his thumb

	Pursuit Aiming	The pursuit aiming test requires subjects to use a pencil to place one dot inside each circle following the pattern given on the printed pursuit aiming test sheet. The task is to do this as quickly as possible for 60 sec
	Santa Ana	The Santa Ana test tests manual dexterity. It requires subjects to remove pegs from holes in a board, turn them 180 degrees and reinsert them. The number rotated correctly within a specified time (30 sec) is taken as the score. It is similar to the Grooved Pegboard test and the Purdue Pegboard test.
Attention/ speed of information processing	Simple reaction time SD	Uses the standard deviation of the simple reaction time test in an individual person as a measure of speed processing.
	Trail making A	Two part test. In part A, the task is to connect consecutively numbered circles by drawing lines between them, while part B requires an alternation between 2 sequences, with numbered and lettered circles, respectively. The time taken to complete the test and the number of errors are recorded.
	Digit symbol	Based on a printed key the subject is asked to combine 9 symbols with the corresponding 9 digits. The raw score is the number of correctly placed symbols in 90 seconds.
Tremor	Static steadiness	Test consist of a metal plate with 9 holes with decreasing diameters from 13 to 2.5 mm. The task is to hold an electric pen in each hole for 15 seconds without touching the metal plate. The measures obtained are the number of touches and touch time. The summation scores for all 9 holes are used.

b. <u>Relationship between demographics and neurological outcomes of manganese exposure</u>

A single review [33]including 19 studies calculated the effect of important covariates (see appendix D) in each of the studies and concluded that all of the cause and effect evidence in primary epidemiological studies is confounded by the variations across the compared groups in important demographic variables. These include but are not limited to co-exposure to other toxic substances (such as lead), variation in age and education levels across groups. When accounted for these variables, appropriately in advance, the relationship between Mn and adverse neurological health measured by nonspecific tests does not remain significant. Authors correctly pointed out that a meta-analysis that ignores the confounding by covariates is of limited use and recommend that primary research should address these limitations by having fewer but more relevant and more valid tests along with better control of demographic variables.

These important variables were included in the set of competitive or confounding factors that must be adjusted for in primary studies to be synthesized in our review for analysing the health effects of exposure.

Respiratory disorders

We included incidence and/or severity (change from baseline) of any diagnosable clinical respiratory condition.[43]

Fertility outcomes

These can range from frequency or number of conceptions to changes in sperm quality or quantity.[43]

Cancer

We included studies reporting incidence or incidence rates.[43]

Neurological disorders

1. Parkinson's disease

Reviews of Mn exposure and Parkinson's disease

A single review (table 5), Mortimer et al. [40] found manganese exposure in welders to be unrelated to or even protective for PD from three included studies.

For the three studies included in this review for this outcome, two were missing entirely from the characteristics of included studies table and the third did not report the number of exposed cases. Of these three studies, one employed self-reported exposure to manganese at work, while the other two relied on job categories as equivalent to manganese exposure. These issues along with the involvement of the welding industry in the review puts into question whether the review's findings are completely free of bias. Thus we believe that the relationship of Mn exposure and Parkinson's disease requires better exploration to come to a clear conclusion.

Table 5 Characteristics of reviews reporting Parkinson's disease as a result of manganese exposure

Review ID	Mortimer, 2012
Aim/ question	To assess association of welding and Mn exposure with Parkinson disease (PD)
Population	Welders/cutters
Exposure	NR, likely welding/Mn as per primary study
Comparison	NR
Primary outcome	Parkinson's disease
Study designs included	Cohort/case-control/mortality studies
Search & selection	Pubmed/CDSR/published reviews
No of studies included	13
No. Included in MA	11 (9 in one and 3 in the other MA)
Primary effect measure	RR
Method of synthesis	Both fixed and random effects MA
Heterogeneity exploration	I^2 statistic, heterogeneity not explored
Publication bias assessment	NR
Results obtained	RR=0.86, 95%CI= 0.8 to 0.92
Authors conclusions	Welding and Mn exposure are not associated with Parkinson's disease
Conflict of interest/ sponsor	First author and the early literature search for this paper was paid towards by welding industry defense group

Primary studies of Mn exposure and Parkinson's disease

Two retrospective cohort studies fitting our inclusion criteria assessed the risk of PD associated with exposure to Mn at work.[49 52] The characteristic of these two studies are tabulated below. The findings were quite heterogeneous (I^2=75%) when pooled. Also, the relatively lower exposure study[52] showing a risk increase (albeit not statistically significant) compared to the other[49] which showed no risk increase.

The studies were both large, and addressed welders with relatively low exposure levels although exposure reporting varied. Forde presented results for the mean exposure level whereas Park presented the risks for each of the three exposure categories, but for all movement disorders of ICD G20-26 together, likely because so few events were found. Both evaluated the same outcomes as defined by ICD codes (G20-G26). However, the studies did not account for all or even the same important confounders.

Table 6 Characteristics of studies reporting Parkinson's disease as a result of manganese exposure

Study ID	Fored 2006	Park 2006 (Korean shipyard study)
Country	Sweden	South Korea
Study design	Retrospective cohort	Retrospective cohort
Duration of follow up years	29 years	NR
Participant source	National census database	Ship building companies
Participants analysed 'n'	Exp=49488; ref=489572; total=539060	Exp= 24,963; ref= 13,597; total= 38560
Mean age years	NR	NR
Gender	Men	Men
Attrition	NR	NR
Occupation/ industry	Welder/any	Welder/ shipbuilding
Exposure data source	Industrial hygiene survey 1974/75	Company records of yearly exposure levels plus job duration
Exposure categories	Mild steel MMA; Mild steel MAG; Railroad MMA; Stainless steel MMA	High, low, very low
Exposure definition	NR	At least one year employment as welder at companies between 1970 and 2000
Air exposure(SD)mean/median/ range mg/m3	MSMMA=0.26; MSMAG=0.3; RRMMA=0.13; SSMMA=0.14	High=0.88 (3.1); Low=0.1 (4); Very low= 0.008
CEI mg/m3. Year	NR	NR
Exposure in the reference/control group	NR	NR
Confounders controlled for	Age; county of residence; education; time period	Age
Outcomes: measure	ICD codes G20 to G26 (PD/basal ganglia disease)	ICD codes G20 to G26 (PD/basal ganglia disease)
Risk estimate	RR (95%CI) PD= 0.89 (0.79-0.99); all basal ganglia movement disorders (including PD)= 0.91 (0.8-1.01)	RR (95% CI) high vs referent=1.96 (0.31-12.5); very low vs referent=3.64 (0.72-18.6); exposed vs referent= 4.2 (0.96- 18.3)

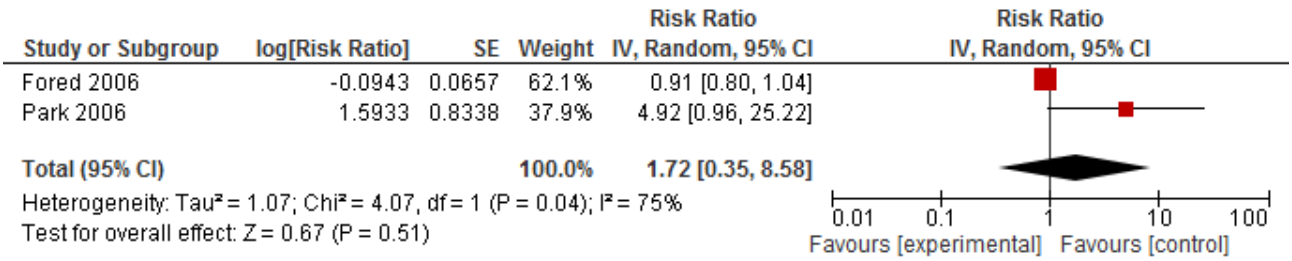

Figure 2: Exposed versus unexposed; ICD G20-26 (PD and other basal ganglia disorders)

Due to a lack of defined exposure categories (upper and lower bounds) we were unable to do a exposure-dose response analysis. Pooling these two studies in a random effects generic inverse variance meta-analysis gives a non-significant increased risk of 1.72 (95%CI= 0.35 to 8.85; I^2=75%) for all exposed compared to all unexposed.

2. Neurological function tests

Reviews of Mn exposure and Neurological function tests

Four reviews [18 35 37 43] assessed outcomes of neurological function using similar tests.

The IEH report[43] was different in that they did not synthesize the evidence from the included 28 studies and instead presented the results of each separately. The studies varied widely in exposure assessment- some had measured exposure in length of time and others in Mn intensity, while some had not measured at all (assumed exposed). Almost all were cross sectional designs, and a wide range of tests was used across the studies which makes pooling difficult. However, no attempt was made to separate well designed studies from the rest. But more importantly, the report did not present any analysis on if and what domains of the neurological function were most related to or affected by the Mn exposure, although this would have made the decision of a no effect level (safe exposure level) easier and more logical. The report concluded that a negative effect of Mn exposure was likely and that limiting the exposure to 0.1 mg/m^3 respirable manganese will prevent most workers from developing the subtlest detectable effect.

The other three included worker populations and the exposure was only occupational. However, the comparison group was not defined per-se in any of these. Included studies tables indicated that the comparison was either workers from a similar work site without exposure to manganese or, in some cases, healthy controls from the adult population.

The review by Ma[37] included a majority of Chinese studies and found neurological tests scores to be correlated with Mn as a negative effect of exposure. However, the effect was not significant ($p>0.05$).

The other two reviews that reported a meta-analysis [18 35] both found a harmful effect of Mn exposure for neurological functions, however, the size and consistency varied. The reviews concluded that the effect was not

enough to confirm a link between Mn exposure and the studied health effects. Only two of these reviews IEH and Meyer-Baron explored heterogeneity in their findings, both finding no significant source.

The Meyer-Baron review of 2009 based on summary data was more explicitly presented than their sensitivity analysis of a subset of studies in an individual patient data meta-analysis reported in 2013[39]. The later analysis found a small harmful effect with more than half of the neurological tests analysed, however, the authors still conclude that a link is unlikely and that the finding is probably due to confounding factors. The review lacks a comprehensive search and a quality assessment in addition to duplicate assessment of studies. These issues again make one question the reliability of the findings.

Primary studies of manganese exposure and neurological function tests

Four longitudinal studies [14 16 44] [54] assessed the previously decided domains of neurological function using the recommended tests.
Characteristics of the studies and the results reported/ calculated from these are presented in the table 7 below:

Table 7 Results of studies reporting neurological test scores as a result of manganese exposure

Study ID		**Blond 2007**	**Bouchard 2007**	**Bowler 2011**	**Roels 1999**
Outcomes measure		Neurological tests (finger tapping; reaction time) CATSYS scores	Neurological tests with computer and questionnaire/interview	Neurological tests with CATSYS and clinical. Change from baseline in exposed only.	Neurological tests PN1, HR, HRR, and VRT
Domain	**Test**	**Results**			
Motor function/Fine motor performance/ Dexterity	**Finger tapping score** High score = better performance	MD = -0.05; 95% CI = -0.24 to 0.14	MD = 1.5	MD 1.5; 95% CI = -1.17 to 4.17	NR
	Reaction time seconds High score = better performance	MD = 0.008; 95% CI = -0.1 to 0.12	NR	NR	Baseline value for VRT: Mn group Mean (SD) = 24.5 (2.1); Control group Mean (SD) = 23.7 (1.8); p = < 0.01
	Luria Motor total score Low score = better performance	NR	MD=1.7	MD =1.86; 95% CI = 1.78 to 1.93	NR
	Pursuit Aiming	NR	NR	NR	PN1 score MD (from fig 4): low exposure group = -1.5; medium exposure group = -1.4; high exposure group= -1.2

	Santa Ana	NR	NR	Baseline values only: Mean = 21.89; SD 3.83; % impaired = 97.7 %	NR
Attention/ speed of information processing	**Reaction time SD**	MD = -0.005; 95% CI = -0.25 to 0.26	NR	NR	NR
	Trail making A Low score = better performance	NR	MD = 4.4	Baseline values only: Mean 44.4; SD = 11.78; % impaired 31%	NR
	Digit symbol High score = better performance	NR	MD = -0.3	MD = 0.76; 95% CI = 0.2 to 1.27	NR
Recognition (visual perception and memory)	**Benton visual retention**	NR	NR	NR	NR
Short term and working memory	**Digit span** High score = better performance	NR	MD = -1	MD = 0.46; 95% CI = 0.04 to 0.87	NR
Tremor	**Tremor intensity** Low score = better performance	MD = -0.005 95% CI = 0.003 to -0.01	NR	MD = 0.24; 95% CI = -0.01 to 0.49	NR
	Hand steadiness time of contacts (sec) Low score = better performance	NR	MD = -0.9	NR	Baseline values H9 score: Mn group Mean (SD) = 16.1 (8.9); Control group Mean (SD) = 13.5 (9.1); p = <0.01
	Hand steadiness number of contacts Low score=better performance	NR	MD = 28	NR	Baseline values HR score: Mn group Mean (SD) = 42.6 (25.1); Control group Mean (SD) = 34.3 (26.2); p = <0.01

The analyses are presented by outcome/ test below.

1) **Motor function/Fine motor performance/Dexterity**

   a. **Finger tapping**

Three studies reported this outcome. The samples were small and studies had a large drop out. Confounding factors were not well adjusted for and because of lack of suitable data for the same, could not be adjusted for the review either. The only study providing enough data for entering into a meta-analysis was Blond (MD = -0.05; 95% CI = -0.24 to 0.14) thus meta-analysis could not be performed. However, the mean difference was not significant. Bouchard

found exposed group score to be 1.5 more than the control group. Bowler reported a before after non-significant improvement in the score of the exposed group after cessation of exposure.

b. Simple reaction time

One study[44] reported reaction time data as mean and standard deviations for both compared groups for baseline as well as end of follow up. Imputing missing data from the provided information, there was no difference between the two groups MD = 0.008 (95% CI = -0.1 to 0.12).

The other study[54] reported only the baseline figures for simple visual reaction time (VRT) and a statement that insignificant time trends were seen with VRT test and exposure.

c. Luria-Nebraska finger-thumb sequential touch

Two studies [14 16]provided a total score for the Luria Nebraska Motor scale. However, data were not pooled due to lack of data on variability. Individually both studies showed better performance for the exposed workers over time (mean change from baseline for exposed in Bowler[14] = 1.86; mean difference between exposed and controls in Bouchard[16] = 1.7).

d. Pursuit Aiming

Roels [54] reported a score of percentange of hits with a stylus on targets on a plate, which can be considered a measure of pursuit aiming. The scores were not reported in figures but a change from baseline for all groups could be obtained from a figure in the report. These indicated a poorer score for all previously exposed groups compared to controls. The mean difference in score was -1.2 for the highest exposure group, -1.4 for medium exposure group and -1.5 for the low exposure group. Measure of variance were not computable from the data available, however we consider the clinical significance of the differences minimal.

e. Santa Ana

Only the baseline study of Bowler 2011 provide the scores for this test for the studied group of 45 welders. The data indicated that impairment based on this test was present in 97% of the welders and the mean scores were 21.89; SD 3.83. However, this outcome was not assessed at follow up so no conclusion can be drawn regarding its relationship with being formerly exposed to manganese. This instance also indicated selective reporting of outcomes.

2) Attention/ speed of information processing

a. Simple reaction time SD

Information was presented regarding this outcome in one study only. The mean difference in the SD of reaction time between the exposed and unexposed suggested that reaction time varies more in the exposed than controls. However, the effect was not significant MD = -0.005; 95% CI = -0.25 to 0.26.

b. Trail making A

Data were presented in two studies for this outcome. Bowler[14]provided only the baseline values for their original sample of 45 welders (Mean 44.4 $\pm$ 11.78; 31% impaired). Bouchard[16]showed a mean difference of 4.40 indicating that unexposed performed better on this test than the exposed, although we do not know of the variability or the significance of this result.

c. Digit symbol

Two studies [14 16]presented this outcome. Bowler found that in previously exposed workers, the performance improved slightly over time for this test (MD = 0.76; 95% CI = 0.2 to 1.27). However, Bouchard found that for change from baseline scores, the controls performed better than the exposed workers (MD = -0.3).

3) Recognition (visual perception and memory)

a. Benton visual retention

No study reported on this outcome

4) Short term and working memory

a. Digit span

The same two studies[14 16] reported this outcome but data were not pooled due to heterogeneity. Bowler found the performance improved over time for prior exposed workers (MD = 0.46; 95% CI = 0.04 to 0.87). Bouchard reported worse performance for change from baseline for the exposed compared to the unexposed (MD= -1).

5) Tremor

a. Static steadiness

Static steadiness (resting tremor) was not reported as such. However, two studies reported some measure of tremor as tremor intensity in m/s^2.[14 44] Bowler[14] reported tremor intensity to improve over time (Cohen's d = 0.24; SD=0.6; p=0.07). Blond[44] reported better performance in exposed also compared to unexposed workers (MD = -0.01).

On the other hand, Bouchard[16] presented hand steadiness as number of contacts and time of contact in seconds. These (presented in table 7 above) show contradictory results: number of contacts improved in the exposed whereas time of contacts worsened. However, the significance is not known.

3. Amyotrophic lateral sclerosis (ALS)

Manganese exposure and Amyotrophic lateral sclerosis (ALS)

The single review addressing this outcome[41] found only one study assessing manganese exposure as a result of pesticide exposure as a risk factor for ALS. This study with few exposed cases had a valid and reliable exposure assessment for manganese and found that the risk of ALS was non-significantly higher for exposed compared to control group. The study findings were based on two exposed and one unexposed participant for manganese exposure where the exposure was assigned by a blinded industrial hygienist using job histories based on 'potential for the exposure of interest'. The study scored less than optimal on the quality criteria applied in the review. Therefore, authors concluded that evidence was not conclusive of a link between exposure of manganese in pesticide and development of ALS and recommended better designed larger studies for all exogenous risk factors for ALS. Their conclusions appear valid in the light of the data. The review was also the highest scoring on AMSTAR (7) which warrants further confidence.

No empirical studies were found that fit our inclusion criteria.

Respiratory disorders

Respiratory disorders have been associated with manganese exposure in cross sectional studies, however, we only included an outcome related to the respiratory system if the incidence or severity as change from baseline were reported. Thus, all valid clinical diagnoses were included if reported as a change from baseline.

Reviews of Mn exposure and respiratory disorders

Only one review addressing the respiratory outcomes was found.[43] Based on eight studies of varied designs this review concluded that there is some evidence of respiratory adverse effects with Mn exposure without linking any specific test or disease to exposure.

Primary studies of Mn exposure and respiratory disorders

Only one study fit the inclusion criteria.[46] This study assessed respiratory conditions of asthma and bronchitis in miners exposed to very high levels of Mn. Although the findings appear valid, these must be considered in context of the high exposure of the cohort. Table 8.

Table 8 characteristics of primary studies of Mn exposure and respiratory disorders

Study ID	Boojar 2002 (Iran miner study)
Country	Iran
Study design	Prospective cohort

Duration of follow up years	7 years likely
Participant source	Iron Mines
Participants analysed 'n'	exp=145; ref=65; total=210
Mean age years	26 (5.3)
Gender	Men
Attrition	NR
Occupation/ industry	Miners/ iron mining
Exposure data source	personal air sampling done for the study
Exposure categories	NR
Exposure definition	Miners; No threshold for CEI reported
Air exposure(SD)mean/median/ range mg/m^3	64 (41)
CEI mg/m^3. Year	at start=1365 (645); at end= 114 (66)
Exposure in the reference/control group	NR; not miner workers
Confounders controlled for	age; salary; education; smoking;
Outcomes: measure	incident occupational asthma/bronchitis= self-report+ clinical diagnosis from medical records;
Risk estimate	RR(CI) bronchitis= 2.35 (0.84-6.58); occupational asthma= 7.7 (0.45- 131.18)
Baseline imbalances	Mn exposure in water (ingestion)

Fertility disorders

These can range from lowered sperm quantity or quality to frequency or number of conceptions or children. Like respiratory disorders, fertility disorders in men have also been reported in literature largely in cross sectional studies. We included studies only if post exposure change from baseline or incident data were reported.

Only one review [43] analysed fertility outcomes of exposure to manganese. Based on two cross sectional studies from Belgium the review concluded that the evidence was conflicting. In addition, that there was a need for better study designs and outcomes more relevant to male fertility than the number of children born to the wife assessed in both studies. No evidence on female fertility was found. We especially could not include studies that used time taken to conceive or time to pregnancy, which is an often used functional measure.

No primary study reporting change from baseline of any male fertility outcomes fit our inclusion criteria.

Cancer

Only incidence of cancer as confirmed from a registry or a biopsy were considered for our review.

Reviews of Mn exposure and cancer

Two reviews assessing the causal link between Mn and cancer were included. [30 43] Assem 2011 was based on the IEH 2004 review and was an update of the cancer component of the report. The review included the same 4 occupational exposure studies as in the IEH 2004 review, all of which suffered from the lack of valid Mn exposure data and involved co-exposure with other potential or known carcinogens. The authors therefore concluded the same as the IEH report, that the evidence to date is insufficient to ascribe any mutagenic or carcinogenic potential to manganese exposure at work. They further recommended that instead of another epidemiological study exploring this link the priority should be given to conducting studies that explore the potential mechanisms of carcinogenicity of manganese which are poorly understood to date.

Primary studies of Mn exposure and cancer

A single primary study that fit our inclusion criteria for this outcome following occupational Mn exposure was Hobbesland 1999.[50] This was one of the studies included in the reviews mentioned above. The study was a retrospective follow up of between 3 to 20 years of a cohort of 6,363 ferro-alloy workers in Norway. Furnace workers with an average exposure level of 0.99 mg/m^3 for at least six months during 1933 to 1991 were compared to non-furnace workers with an assumed zero exposure during the same time period. Additional exposures in furnace workers that may be carcinogenic were also not accounted for. The outcome of cancer was obtained from the national cancer registry for the time period between 1951 and 1991. Adjusting only for the confounder of employment duration the study found an SIR of 1.02 (95%CI=0.94 to 1.10), even though some specific cancer sites showed an increased risk. The reported exposure level was a median value from limited personal sampling in one of the factories in 1991 only. In addition, there could be confounding by exposure to substances other than manganese in these work environments that can be carcinogenic, which have not been considered. See table 9.

Table 9 characteristics of primary studies of Mn exposure and cancer incidence

Study ID	Hobbesland 1999 (cancer risk in welders study)
Country	Norway
Study design	Retrospective cohort
Duration of follow up years	3 – 20 years
Participant source	Ferro alloy plants
Participants analysed 'n'	Total = 6363
Mean age years	NR
Gender	Men
Attrition	NR
Occupation/ industry	Furnace workers / steel industry
Exposure data source	Personal samples likely from company records

Exposure categories	Furnace workers, non-furnace workers
Exposure definition	Six months employment at any of the four plants between 1933 to 1991
Air exposure(SD)mean/median/ range mg/m3	0.99
CEI mg/m3. Year	NR
Exposure in the reference/control group	NR
Confounders controlled for	Employment duration
Outcomes: measure	Cancer incidence as per national database
Risk estimate	SIR for furnace workers = 1.02 (0.94 - 1.10)

Grading of evidence

- Danish Occupational Medicine Association Grading

According to this standard, the quality of evidence for the causal link between Mn exposure and any of its health effects is insufficient (0). This means that the available studies are of insufficient quality, consistency, or statistical power to permit a conclusion regarding the presence or absence of a causal association.

- GRADE evidence level

As stated a-priori, we downgraded or upgraded evidence for each outcome across studies (see table 10). We found that for most outcomes the evidence is of very low quality.

Tabel 10 GRADE evidence rating for each health outcome attributed to Mn exposure

Outcome	Downgrading					Upgrading			Quality of evidence
	Risk of bias/ design limits	Inconsistency	Indirectness	Imprecision	Publication Bias	Large effect size	Confounding does not reduce effect	Dose Response	
Parkinson's disease	-1	-1							Very low quality
Short term and working memory	-1	-1	-1		-1				Very low quality
Motor function/ Fine motor performance/ Dexterity	-1	-1	-1		-1				Very low quality
Attention/ speed of information processing	-1	-1	-1		-1				Very low quality
Tremor	-1	-1	-1		-1				Very low quality
Cancer	-1	-1							Very low quality
Occupational asthma/ Bronchitis	-1	-1			-1				Very low quality

Discussion

Summary of main findings

We found very low quality evidence indicating uncertainty ranging from no effect to a very large effect of Mn exposure on incidence of Parkinson's disease. Results in the two significantly heterogeneous studies, which appear to show an excess risk were not statistically significant and the confidence intervals are wide. The exposure did not affect finger tapping scores considerably based on data from three studies.

For other neurological test outcomes, we found very low quality evidence of inconsistent effects of Mn in between one to three studies.

For respiratory disease and cancer there was very low quality evidence from a single study that Mn did not lead to any significant increased risk for either type of diseases in the exposed.

No evidence was found on fertility problems in men.

Overall completeness and applicability of findings

The range and number of reviews included was wide and varied. However, only two were satisfactory in terms of quality on the AMSTAR. Thus, the larger set of reviews and meta-analyses provided very little input to our final results and conclusions, although they indicate more areas for future research. Contrarily, the range of the empirical studies is small and the number of studies providing data for any one outcome is even smaller. Nearly half of these are from northern Europe and therefore the findings do apply to these populations.

Nature of Exposure

The review findings indicated that Mn makes up only about 4% of the welding fumes but we cannot be sure about it because of no accounting for the variation within the welding types. Furthermore, the findings come from a single review of low quality (AMSTAR score is 2). Therefore, we are unable to say with certainty what percentage of welding fumes contain Mn. The Mn in current welding exposures can be determined from valid personal measures for both total welding fume and its corresponding respirable Mn component. At least in some parts of the world these values are recorded annually as part of industrial hygiene requirements. It is important, however, that such a study separates the exposure levels for the different welding types. An average value is only useful if we know what type of welding and welding environment it refers to, considering some types may involve a zero concentration of manganese

and some much more than the average. The individual welder exerts also a considerable influence on the fume generated in some welding processes and on his or her resulting exposure.

The correlation between Mn in air and Mn in the body remained unclear due to inconsistency in findings across the reviews. Although the findings of Li 2014[36] indicate that the Pallidal Index is a reliable measure of Mn burden in the brain, these findings should be interpreted with caution as the authors themselves advise. It still needs to be evaluated how the PI values correlate with valid tests of neurological disease. Evidence on blood Mn, as assessed by two reviews, is also inconsistent, where one review [18]found the measure to be of no use as a biomarker of exposure, while the other, [32]concluded that Mn in blood is an effective biomarker for air exposures above 10 $\mu g/m^3$ (0.001 mg/m^3). This exposure level may exist in many work places, even in workers assumed to be unexposed.

The rationale then, for assessing biomarkers can be that the biomarkers are considered a proxy of *both* the exposure and the outcome/adverse health effect/disease). The next logical question is whether these biomarkers, which are used as proxy for exposure, correlate consistently with the adverse health effect. Because without a clear relationship between these three it would be difficult to attribute the adverse health effects to Mn in air, invalidating the use of the biomarker as a proxy for either the exposure or the outcome. The only review that addressed it was Meyer-Baron 2009 where they explored a relationship between MnB and many neurological tests (only finger tapping, digit span presented) concluding that a clear relationship between the internal exposure biomarker of MnB and effect did not exist.

Based on the low quality and variable methods of all the reviews included for this aspect of our review, we find that, at the moment, the evidence is inconclusive for what constitutes a control group without exposure, or a NOAEL or the use of Mn in blood or urine as measures of exposure.

Adverse effects of Mn exposure

Parkinson's disease (PD)

In the absence of duplicate inclusion or a quality assessment of the studies in the Mortimer review, we considered their findings of no link between Mn exposure and PD at a high risk of bias. Furthermore, information on why studies were excluded is not available, for example, in this review, a prospective cohort with actual Mn measurements in a welding cohort was excluded but a reason was not provided. The findings of our meta-analysis of two follow up studies showed a non-significant 72% increase in risk (RR=1.72; 95% CI= 0.35 to 8.85; I^2=75%). However, given the different results in each study, this is not a convincing result. At best, these results can be attributed to welders with an exposure ranging from 0.008 to 0.26 mg/m^3 of Mn in air, for an unknown particle size or exposure duration. The limitations of the data quantity and quality prevent any clear conclusions.

Neurological function tests

The four reviews individually found conflicting results, with no review clearly stating a conclusion of a confirmed negative effect of Mn exposure at work. Neither did they provide a reliable conclusion on which neurological functions were most or earliest affected by Mn exposure, and to what extent. The variation in the review designs and methods precluded combining of their findings by us. Therefore, we restricted our conclusions to the analysis of the included empirical studies in the following section.

With the exception of one [44]enough data were not available for making meta-analysis possible for any outcomes from the four primary studies assessing neurological test scores. However, the mean differences reported by Blond in finger tapping, simple reaction time, simple reaction time SD, and tremor intensity were all clinically very small and statistically non-significant ($p > 0.05$). Bouchard[16] provided the mean differences for most of the tests we defined a-priori for this review. However, in the absence of variability measures we were unable to analyse these together with Blond.[44] Authors were sent a request for data and a response is awaited.

Bowler[14] and Roels[54] reported change as a result of reduction of exposure and in both studies an improvement is seen in neurological test scores after 15 and 10 years of follow up, although only limited data were reported (missing final scores and standard deviations) and there were signs of selective outcome reporting in both studies. Also, Bowler[14] did not compare these findings with a control group while the control group used by Roels[54] had significantly better test scores at baseline. This means that although an improvement in the mean values is there, it is difficult to ascribe these to the decrease in Mn exposure.

In summary, we are no closer than the previous reviews in terms of quantifying the link between Mn exposure and neurological test score changes, although our findings indicate areas for focus in the future empirical studies.

This should be viewed also in context of the fact that these tests on their own are not diagnostic of any preclinical neurological condition or even for progress of an existing condition: a combination of clinical exam with these a range of these test is needed for a diagnosis for a preclinical disease also. There are also variations within normal ranges of these tests within the same population over time due to practice or fatigue, which can be in either direction. There is also evidence from previous research that specificity of these tests is poor and these often, when used alone, lead to misclassification of a normal subject as with neurological damage.[62] Unless and until normative ranges are available for healthy populations, with research based cut off for thresholds of abnormal values, it may be inappropriate to rely on these tests since differences have been shown even among healthy control groups. [63 64]

Respiratory and fertility outcomes

With only one review and one study, very limited data is available for respiratory conditions. There seems to be an increase in risk of occupational asthma and bronchitis in Iranian miners but additional exposures cannot be ruled out, besides the very high level of Mn reported for this study, which is rare today. Dust of other types along with many other factors can lead to occupational asthma c.[65] A well-controlled future study can answer this question better.

Fertility outcomes were assessed in one review only and we did not find any primary studies evaluating any of the fertility outcomes prospectively. This means that, for now, there is a lack of evidence on the question, but it should be noted that this is not the same as evidence of no effect of Mn on fertility.

Cancer

With two reviews and one primary study, we found that there is no significant rise in risk for cancers in Mn exposed workers, however, the low quality of both the reviews and that of the primary study decrease confidence in this finding. In the absence of clearly defined theory and mechanism of cancer generation that can be attributed to Mn[30], it is best to focus on more basic research that can confirm whether Mn is capable of causing cancer in humans in the first place.

Quality of the evidence

In terms of methodological quality, only two of the reviews scored more than 5 on the AMSTAR.[36 41] Both of these were funded by Academic/ government grants, and declared that none of the authors had any conflicts of interests. Both carried out a comprehensive search and were the only two reviews that indicated duplicate study selection and data extraction. Primary studies also performed poorly on the quality criteria, with none achieving a low risk of bias judgement overall. Although outcome measurement in majority of the primary studies was found at low risk of bias. This is good because it means the outcome assessment is an area where little improvement is needed and the results can be trusted from the studies at low risk. However, of the low risk studies three used components of a computerized system for neurological tests scores (CATSYS). What we do not know is how valid an individual component test score is for the neurological domain it measures. Evidence exists to indicate that some of its components may not correlate at all with the symptoms of PD.[66] The use of a validated scale in full is best, for example the Unified Parkinson's Disease Rating Scale (UPDRS) for diagnosing PD even though it requires a neurologist. It must be noted, however, that the UPDRS is a scale to characterize and assess severity of a clinical picture. It cannot discriminate manganism, Parkinsonism and secondary PD from primary idiopathic PD.

We stated a priori that evidence would be considered high quality if it were from a well-controlled prospective observational study. Therefore we limited our findings to those reviews and studies that used suitable design (change from baseline for an outcome) when examining causal effect of Mn in air on neuropsychological health. Based on our assessment all evidence is at best of low to very low quality and therefore insufficient for causality.

Considering the evidence on cancer, and respiratory disorders the quantity and quality was found very limited and so insufficient for causality, whereas there was no prospectively collected evidence was found on outcomes of fertility.

Thus, the evidence within this overview is overall insufficient for causality of any adverse health outcome, based on the scientific criteria of the Danish Occupational Medicine Association and of very low quality based on GRADE.

This prevented us from making any clear claims on the nature and extent of the relationship between occupational Mn exposure, its biomarkers, and any health effects attributable to this exposure.

Potential biases in the process of the review

Exclusion of cross sectional study data

We excluded all studies providing an assessment of exposure and outcome at one point in time only, where an exposure measurement at start of follow up and a change from a baseline measurement in the exposed group were not reported. This excluded the majority of the research base that exists on this topic. We consider this exclusion necessary if we are to establish the minimal exposure doses of Mn in air that can cause any harm, and protect the workers in future. Mn exposures at work have consistently declined in the past couple of decades and are considerably lower today than in the 1980s, for example. A study that measured exposure intensity in 2000 for example and correlated this exposure level (either alone or in a composite index) with the health status in the same year (as measured by a test of neuro-motor function) in fact disregards the higher levels of exposure that would have existed for the cohort under study in the previous two decades. Thus, unless an accurate exposure estimation for the entire duration, but especially that prior to appearance of the outcome of interest is provided, an analysis will potentially underestimate the effect of the exposure.

We had a closer look at 143 studies that we excluded at the title-abstract stage because of the cross-sectional design. Of these 70 had studied some kind of neurological or neuropsychological outcome but many other outcomes were studiedsuch as erythrocyte level, immunological outcomes and other proteins in blood. Many articles were by the same author or the same group of authors reporting each outcome measure separately such as 18 articles by Misiewicz. Most articles (N=62) were a comparison of welders versus non-welders without properly reporting Mn exposure. The oldest study was from 1965 and the number of studies during the seventies was 7 and during the eighties 11. After that, there is a steep increase with 46 studies between 1990 and 2000, 48 studies between 2000 and 2010, and already 31 studies thereafter. It is difficult to understand why so many cross-sectional studies are still reported other than that it is an easy way to publish an article.

Exclusion of haematological outcomes

We exclude haematological outcomes such as Mn levels in blood and cell count deficiencies in the human blood. This was because the blood composition is dynamic and is affected by many factors unrelated to the exposure also, for example iron metabolism. The confounding effects of such factors are rarely accounted for and are sometimes not well documented either. Thus, these outcome measures provide unreliable evidence for either prevention or causal inference.

Language and publication bias

We incidentally came across a report [67] outside of our extensive search which contained reference to a study in the grey literature also not found by our search.[68] We were unable to obtain a full text of this primary study to enable inclusion till the date of submission of this report. Although our conclusion do not seem to differ from the study's abstract, the finding of this study itself indicates that relevant primary literature on the topic may exist outside of the peer reviewed and even some grey literature databases. In addition, in the current review, contact with all primary study authors could not be established, largely due to time constraints. This means that the findings of our review refer to the published component of the included studies and this should be kept in mind when interpreting our findings. As indicated by finding several reports of studies that were not published in an official scientific journal, we believe that a language and publication bias may exist in this field without accounting for which we cannot claim the evidence to be complete. Given that both reports were on studies that had not found evidence for an health effect of Manganese, we suppose that there is a bias towards positive findings in the official literature. Future systematic reviews need to take into account the issues as highlighted here.

We found a Chinese language review also, which included at least 14 Chinese language studies of what appeared to be short-term exposure. Although we could get this review translated and included, we could not locate full texts of their included Chinese studies for assessment due to lack of time and resources. How many more such studies in various languages remain outside the mainstream English language circle can only be imagined. China is the largest producer of manganese alloy, as well as a major contributor to global reserves of the Mn ore along with South Africa, India and Gabon. [69] Mining and milling manganese ore may still be a major source of exposure in these countries, as indicated in the Boojar 2002[46] study on miners.

Conclusion

We found insufficient good quality data to refute or assert a causal association between recent or current levels of occupational exposure to manganese and adverse health outcomes. We could not assess an exposure-dose repose association because of lack of suitable data. We could not find a threshold above which harmful effects start appearing because these varied widely, with inconsistency in findings. In the absence of reliable and consistent data, we conclude that the causal link between adverse health effects of occupational manganese exposure currently lacks enough good evidence. We find that for most neurological outcomes, exposure to manganese cannot be confirmed as a causal factor due to poor exposure assessments, confounding, large losses to follow up and incomplete presentation of outcomes. This does not mean that a link does not exist, but only that it is not apparent in the evidence we examined.

The current exposure limits have been derived from cross sectional or data linkage studies that assessed neurological function tests. The studies used self-reported exposure length for the past or a current single measurement of intensity, which was then matched with the current neurological function test status to provide a correlation between the prevalence of exposure and effect. The occupational exposure limits are then formulated conservatively from these estimates to ensure prevention of harm to workers. Therefore, although these studies provide sufficiently suitable data for preventive efforts, these do not provide enough evidence for establishing causation. What these studies may also

hide is a possibility that the actual NOAEL values could be even lower than estimated, because of the poor quality assessment of exposure effect relationship.

Implications for research

An ideal study examining the effect of manganese exposure could be based on up to date industrial hygiene data on exposure of workers, if available. What, however, is needed most is a careful a-priori set of hypotheses for exposure thresholds which are to be tested.

A consensus of the leaders in the field of Mn exposure is needed for deciding what constitutes exposure and non-exposure at work, and what should be a composite index of exposure and how it should be measured. This should also take into account what the influence is of measurement strategies such as group-based or individual based. For example, the ambient levels allowed for metropolitan cities for manganese could be considered the threshold for control or the unexposed. Then, any exposure at work place above this level could be agreed upon as occupationally relevant exposure. Above this threshold, exposure could be designated as low medium and high exposure. For example, low exposure groups could be those working in areas close to the furnace rooms at an alloy plant. Mining exposure particles are different from welding exposure particles so a different set of thresholds for each, or defining the most critical particle size with respect to health would be needed. This would require measuring air manganese levels precisely at different work areas and at different times. The comparability of outcome measures in future studies will enable combining of their results, which is the advantage that meta-analysis offers over small independent studies. A consensus of leaders in the field is needed to enumerate the valid and clinically relevant outcomes of occupational Mn exposure.

The various pathophysiological stages on the path from manganese exposure to health effects are still unclear. These need elaboration by in-vitro and in-vivo research to generate clear hypotheses, before any further resources are engaged in assessing a presumed effect in humans.

Appendices

Appendix A: search strategy

Search	Add to builder	Query	Items found	Time
#8	Add	Search **(((((((("Manganese"[Mesh] OR "Manganese Compounds"[Mesh] OR manganese[tw]))) OR (("Manganese Poisoning"[Mesh] OR manganism* OR "Manganese/adverse effects"[Mesh] OR "Manganese/toxicity"[Mesh] OR "Manganese Compounds/adverse effects"[Mesh] OR "Manganese Compounds/poisoning"[Mesh] OR "Manganese Compounds/toxicity"[Mesh]))))) AND (((work[tw] OR works*[tw] OR work'*[tw] OR worka*[tw] OR worke*[tw] OR workg*[tw] OR worki*[tw] OR workl*[tw] OR workp*[tw] OR occupation*[tw]) OR welder OR welders OR welding OR "steel industry" OR mining OR miners)))) NOT ((Animals[Mesh] NOT Humans[Mesh]))**	2272	03:16:10
#7	Add	Search **(Animals[Mesh] NOT Humans[Mesh])**	3875032	03:15:47
#6	Add	Search **((((((("Manganese"[Mesh] OR "Manganese Compounds"[Mesh] OR manganese[tw]))) OR (("Manganese Poisoning"[Mesh] OR manganism* OR "Manganese/adverse effects"[Mesh] OR "Manganese/toxicity"[Mesh] OR "Manganese Compounds/adverse effects"[Mesh] OR "Manganese Compounds/poisoning"[Mesh] OR "Manganese Compounds/toxicity"[Mesh]))))) AND (((work[tw] OR works*[tw] OR work'*[tw] OR worka*[tw] OR worke*[tw] OR workg*[tw] OR worki*[tw] OR workl*[tw] OR workp*[tw] OR occupation*[tw]) OR welder OR welders OR welding OR "steel industry" OR mining OR miners))**	2671	03:15:30
#5	Add	Search **((work[tw] OR works*[tw] OR work'*[tw] OR worka*[tw] OR worke*[tw] OR workg*[tw] OR worki*[tw] OR workl*[tw] OR workp*[tw] OR occupation*[tw]) OR welder OR welders OR welding OR "steel industry" OR mining OR miners)**	1118429	03:15:16

Search	Add to builder	Query	Items found	Time
#4	Add	Search **((((((("Manganese"[Mesh] OR "Manganese Compounds"[Mesh] OR manganese[tw])))) OR (("Manganese Poisoning"[Mesh] OR manganism* OR "Manganese/adverse effects"[Mesh] OR "Manganese/toxicity"[Mesh] OR "Manganese Compounds/adverse effects"[Mesh] OR "Manganese Compounds/poisoning"[Mesh] OR "Manganese Compounds/toxicity"[Mesh]))))) NOT ((("Manganese"[Mesh] OR "Manganese Compounds"[Mesh] OR manganese[tw]))**	3	03:14:20
#3	Add	Search **((((("Manganese"[Mesh] OR "Manganese Compounds"[Mesh] OR manganese[tw])))) OR (("Manganese Poisoning"[Mesh] OR manganism* OR "Manganese/adverse effects"[Mesh] OR "Manganese/toxicity"[Mesh] OR "Manganese Compounds/adverse effects"[Mesh] OR "Manganese Compounds/poisoning"[Mesh] OR "Manganese Compounds/toxicity"[Mesh]))**	37606	03:13:52
#2	Add	Search **("Manganese Poisoning"[Mesh] OR manganism* OR "Manganese/adverse effects"[Mesh] OR "Manganese/toxicity"[Mesh] OR "Manganese Compounds/adverse effects"[Mesh] OR "Manganese Compounds/poisoning"[Mesh] OR "Manganese Compounds/toxicity"[Mesh])**	1784	03:13:43
#1	Add	Search **(("Manganese"[Mesh] OR "Manganese Compounds"[Mesh] OR manganese[tw])**	37603	03:13:13

Appendix B: data extraction form

Data extraction form for systematic reviews and meta-analyses

Review ID	
Population	
Exposure	
Comparison	
Outcome	
Study designs included	
Search & selection	
No of studies included	
No. Included in MA	
Primary effect measure	
Method of synthesis	
Heterogeneity exploration	
Publication bias assessment	
Results obtained	
Authors conclusions	
Conflict of interest/ sponsor	
AMSTAR rating total	
1. Was an 'a priori' design provided?	
2. Was there duplicate study selection and data extraction?	
3. Was a comprehensive literature search performed?	
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?	
5. Was a list of studies (included and excluded) provided?	
6. Were the characteristics of the included studies provided?	
7. Was the scientific quality of the included studies assessed and documented?	
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	
9. Were the methods used to combine the findings of studies appropriate?	
10. Was the likelihood of publication bias assessed?	
11. Was the conflict of interest included?	

Data Extraction Form Primary Study

Review of effects of occupational exposure to manganese

Study ID (author year)

Notes:

1. Please state **NR** in the response column if an item is not reported in the study
2. Adjusted values are preferred when provided compared to crude ones. if unable to judge please state next to the values or item, for example, 'not clear if adjusted'

3. Since data would be available in many formats a general rule is: when in doubt, take out as much data or information as possible. If a column does not seem to fit the data provided in the paper please describe in the authors words or your own with actual values
4. In the risk of bias assessment the judgement boxes☐ are supplemented with description of situations where that judgement would apply. Also, text boxes are available next to high and low risk judgements for quotes from the study or your comments that made the decision possible. Any explanations would enable quick agreements possible and are encouraged.
5. Please cite the references(author 1, title, journal, year, volume and pages) to other potentially relevant studies cited in this included study here :
6. Any additional report(s) of the same study used /to be used for data extraction (Author 1, title, journal, year, volume, pages) should be indicated here:
7. Any info not available in the paper that is needed from authors should be cited here:

PLEASE SEND THE COMPLETED FORM BACK TO sharea.ijaz@ttl.fi

Your name Date:

Study Characteristics		
Aim/ hypothesis of the study		
Study design (mark at least one that best describes and any other that may apply)	☐Cohort (prospective) study with concurrent controls ☐Cohort (retrospective) study with concurrent controls ☐Case-controlled (retrospective) study ☐Cohort (prospective) study with historical controls Mark if the study had defined populations that were prospectively followed in an attempt to determine distinguishing population characteristics with historical controls ☐Nested case-control Mark if the study started with the identification of persons with a disease of interest and a control (comparison, referent) group without the disease that were identified within the cohort of the subjects, participants in prospective cohort study. The relationship of an attribute to the disease was examined by comparing diseased and non-diseased persons with regard to the frequency or levels of the attribute in each group. ☐Other-specify Specify reported study design with terminology different from the definitions of the National Library of Medicine (described above)	
Study Country		
Participants		
	Exposed	Unexposed
Sources of participants		
Number of Participants (enrolled)		
Number of Participants (analysed)		
Age range or Mean +SD in yrs (describe if reported in another way)		
Gender (% if M+F)		
Years of employment		
Occupation		
Industry		
Attrition rate		
Excluded from analysis		
Notes/ miscellaneous		
Exposure Information		
Source of information on exposure (Interviews, registers etc)	Exposed	Unexposed
Exposure definition		
Exposure measurement		
Unexposed/Reference Category definition		
Exposure categories reported	category name and / or description	

	exposed controls				
Exposure type (mark all that apply)	Dust ☐ fumes☐ blood conc ☐ ambient conc ☐ Other☐ (describe) Not reported☐				
Average exposure duration (yrs m ± sd)	Exposed Unexposed				
Average exposure intensity (particle density/ m3 of air)	Exposed Unexposed				
Average Cumulative exposure (intensity x duration)	Exposed Unexposed				
Other measure of exposure reported (e.g. particle size)	exposure measure name exposed controls				
Confounders					
Confounding factors controlled for	alcohol intake ☐ Iron levels ☐ liver status ☐ education ☐ socioeconomic status ☐ smoking ☐ age ☐ sex ☐ ☐other				
Outcome					
Outcome(s) names, definitions, measures	name	definition	measure		
Outcome Mean (SD) change from baseline	Outcome name	Exposed (Mean (SD) n=	Control Mean (SD) n=	Other measures e.g. p value	MD (SE)
Outcome proportion of people with health effect (N)	N in Exposed	N in Unexposed	Total exposed Total unexposed		RR (CI)
Continuous outcomes					
risk per year of exposure increase	Crude RR 95%CI Adjusted RR 95%CI				
categories of exposure duration	category name	RR/ MD	95% CI		
	cat				

	cat			
	cat			
	cat			
risk per unit of intensity increase	Crude RR 95%CI		Adjusted RR 95%CI	
categories of intensity	category name	RR	95% CI	
	cat			
	cat			
	cat			
	cat			
	cat			
risk per year of cumulative dose increase (int x year or any other)	Crude OR 95%CI		Adjusted OR 95%CI	
categories of cumulative dose	category name	RR	95% CI	
	cat			
	cat			
	cat			
	cat			
	cat			
risk with other exposure measure/unit increase	Crude RR 95%CI		Adjusted RR 95%CI	
categories of other exposure	category name	RR	95% CI	
	cat			
	cat			
	cat			
	cat			
	cat			
Any other outcome	name	RR	95%CI	

Methodological Evaluation of Observational Research (MEVORECH) – observational studies of risk factors of chronic diseases. Adapted from Shamliyan et al. 2010 for Mn exposure and adverse health outcomes review

(as a general rule low risk means best practice, major flaw is high risk, and minor flaw is moderate risk as agreed upon between two independent assessors, poor reporting means unclear risk)

Study ID; Author+ year:
Reviewer ID:

Study characteristics:
Location- country or countries of the residency of the subjects of the study:

Journal of publication:
Year of publication:

Funding source of study:

Responses	Instructions/quote comment	Judgment
Not reported		Poor reporting/ unclear risk
Industry (one or more corporate sponsors)		High risk
Grant(one or more public/not-for-profit sponsors)		Low risk
Combined industry + Grant		Moderate risk
Other - specify	Abstract funding sources that can be specified as corporate or not-for-profit organizations.	

Role of funding organization in data analysis and interpretations of the results

Responses	Instructions/quote comment	Judgment
Not reported		Poor reporting
Sponsoring organization participated in data analyses		High risk
No role of the sponsor organization in the study conduct and reporting		Low risk
Other (specify)	Abstract relevant information about participation of the sponsor in data analysis	

Conflict of interest:

Responses	Instructions/quote comment	Judgment
Disclosure not reported		Poor reporting
Reported not having conflict of interest		Low risk
Reported having conflict of interest (at least one author)		High risk
Other		

Ethical approval of the study:

Responses	Instructions/quote/ comment	Judgment
Not reported		Poor reporting
Study was approved by Ethical Committee		Low risk
Other (specify)	Abstract relevant information whether the study complied with ethical principles of research	

Aim of the study

Response	Quote /comment	Judgment
Aim was not stated		Poor reporting
Aim included association with risk factors without clear definition of the target population		Minor flaw- moderate risk
Aim included association with risk		Low risk

factors with clear definition of the target population		

Internal Validity

Domain 1: Adequate diagnosis of *Outcome*

Source of dependent variable data (target/outcome)

Responses/ options	**Quotes/ comments**	**Judgment** low/unclear/high risk of bias
Objective validated diagnostic methods for the purpose of the study (independent of health care records) e.g. PET OR clinician diagnosed OR confirmed with incidence data from a registry based on valid clinical diagnoses		Low risk
Not reported		Poor reporting- unclear risk
Proxy reported		Major flaw- high risk
Obtained from medical records or administrative databases (mining of data collected for health care purposes)		Minor flaw- moderate risk

Severity, degree of the symptoms of the condition

Severity can be relevant but not assessed in the study		High risk
Severity assessed appropriately		Low risk

Validation of outcomes measurements

Authors report validation of methods of measurement for the study		Low risk
No information about validation		Poor reporting- unclear risk
The authors did not validate the methods to measure dependent variables (no valid methods were obtained)		Major flaw- high risk

Domain 2: Adequate ascertainment of *Exposure*

Source of exposure data

Responses	**Quote/comment**	**Judgment**
Not reported		Poor reporting- unclear risk
Obtained for study from **participants** by self-report/ interviews		
Proxy collected for the study (job matrix, job title records)		High risk
Obtained from **medical records/ administrative database** (mining of data collected for health care purposes)		High risk
Employers' prospectively collected database OR employees prospectively recorded data (e.g. Mean Mn levels over the year in an area)		Moderate risk
Directly measured quantities of **Mn in environment** for the purpose of the study (ambient levels, personal dosimeters)		Low risk of bias
ANY OTHER		

Definition of the exposure/case definition - general

Responses	**Quote/comment**	**Judgment**
Not reported		Poor reporting- unclear risk
Authors define and quantify the exposure for a certain job/occupation		Low risk of bias
Definition of exposure/ case is categorical with an arbitrary threshold (1 yr or more in a job with Mn exposure) OR is		High risk of bias

irrespective of variation in jobs/occupations		
ANY OTHER		

Definition of the exposure/case definition- Reference period/length of exposure

Responses	Quote/comment	Judgment
Reference period/length of exposure not reported		Poor reporting- unclear risk
Reference period/length of exposure not included in definition of the exposure		High risk
Reference period/length of exposure is included in definition, is different from recommended but justified		Low risk of bias
Reference period/length of exposure is included in definition but different from recommended and not justified		High risk
Reference period/length of exposure is included in definition according to consensus/ guidelines		Low risk of bias

Definition of the exposure/case definition - Intensity/dose of exposure

Responses	Quote/comment	Judgment
Intensity/dose is not reported		Poor reporting
Intensity/dose not assessed in the study		High risk
Intensity/dose of exposure as stated by consensus/guidelines included in the definition/assessment of exposure.		Low risk of bias

Measurement of exposure (when not reported -assumed and implied exposure- study will be excluded)

Method used	Quote/ comment	Judgment
The authors reported inter-method validation for the study (one method vs. Another)		Minor flaw-- moderate risk
The authors did not use validated the methods to measure exposure		High risk of bias
Ambient exposure levels or inhalable levels alone		Minor flaw-- moderate risk
Exposure levels measured as ambient/inhalable levels, years of exposure and particle sizes either in a composite measure or separately		Low risk of bias
Any other		

Masking of exposure status for investigators who measured dependent variables (outcomes) and vice versa

Responses	Quote/comment	Judgment
Not reported		Poor reporting
Was possible but not obtained		Minor flaw- moderate risk
Assessors of exposure (Mn) did not know outcome status (e.g. Neurological deficit level) / vice versa		Low risk of bias

Reliability of exposure estimates

Responses	Quote/ comment	Judgment
Not reported		Poor reporting
Intra-observer variability is reported with subjective judgment of reliability		Minor flaw- moderate risk
Good inter observer reliability achieved/ objective measure used/ not applicable for the measure used		Low risk of bias

For case-control studies

Responses	Quote/ comment	Judgment
The authors did not state that the same methods were used to measure exposure risk factors, independent variable) in cases and controls		Minor flaw- moderate risk
The authors used different methods to		Major flaw- high risk

measure exposure in cases and controls		
The authors used same methods for cases and controls to measure exposure		Low risk of bias

Domain 3: Confounding factors bias - inclusion

Responses	Quote/comment	Judgment
Not reported		Poor reporting
Major confounding factors/effect modifiers (magnitude of confounding 30% or more) were not assessed		Major flaw- high risk
Major confounding factors /effect modifiers were assessed partially/ inadequately		Minor flaw-
Major confounding factors/effect modifiers were assessed (known sets of confounders specific for research questions)		Best practice- low risk

Confounding factors - measurement

Responses	Quote/comment	Judgment
Not reported		Poor reporting
Unknown validity to measure confounding factors		Minor flaw- moderate risk
Non-valid methods to measure confounding factors		Major flaw- high risk
Confounders measured with valid methods		Low risk of bias

Domain 4: Attrition bias
For cohort study
Loss of follow-up (acceptable important cut off =)

Responses	Quote/comment	Judgment
Not reported		Poor reporting
Loss of follow up is larger than acceptable		High risk
Loss is minimal/ acceptable		Low risk

For case-control studies

Responses	Quote/comment	Judgment
Not reported		Poor reporting
% of nonresponse differed among cases and controls		Minor flaw- moderate risk
% of nonresponse reported for cases only		Minor flaw- moderate risk
% non-response was reported for both cases and controls and did not differ		Low risk of bias

Domain 5: Analysis of the study
Appropriateness of statistical model to reduce research specific bias

Responses	Quote/comment	Judgment
Strategies/ statistical methods to reduce research specific bias not reported		Poor reporting
Authors did not use statistical models that may be appropriate according to the published literature (examples may include population stratification bias in case-control studies of genetic association, using odds ratio in cohort studies of common diseases, not accounting for missing data or large loss of follow-up)		Major flaw-high risk
Authors did not justify their choice of statistical models to reduce research specific bias		Minor flaw- moderate risk
Authors attempted to reduce bias in post hoc statistical adjustment		Minor flaw- moderate risk
Authors reported more than one method to reduce bias		Lowest risk of bias

Methods reported - tick all applicable		Instructions	Judgment
1	Standardization	Mark if the study controlled for confounding the analysis applying of weighted averages of the relevant measures of disease frequency	

2	Matching	Mark if the study selected the subjects to provide equal distribution of confounding factors in comparison groups	
3	Adjustment in multivariate model		
4	Stratification	Mark if the study controlled for confounding by separating a sample into several subsamples according to specific criteria; therefore the association between exposure and outcomes was evaluated within homogenous categories of confounder.	
5	Propensity scoring	Mark if the study obtained propensity scoring method creating subgroups of exposed and unexposed subjects to be comparable with respect to their distributions of observed confounding factors	
The authors did not obtain methods to reduce bias		Mark if the study did not obtain statistical methods to reduce bias	High risk of bias
Several methods to reduce bias		Mark if the study obtained more than one method to reduce bias	Low risk of bias
Only one method used		Mark if the study used only one method and probably not sufficiently controlled bias	Moderate risk

Dose response with exposure

Responses	Quote/ comment	Judgment
Not assessed		Minor flaw- moderate risk
Not reported		Poor reporting
Dose response assessed in analysis		Low risk of bias

Reporting of tested hypothesis

Responses	Quote/ comment	Judgment
Unclear reporting of the estimates (unclear model, reference level, set of confounding factors…)		Poor reporting
Crude estimates presented only		Major flaw- high risk
Incomplete/ selective reporting of the tested hypotheses (compared to aim and objectives)		Minor flaw-- moderate risk
Adjusted estimates presented for all hypothesis tested as per aims		Low risk of bias

Precision of the estimates

Responses	Quote/ comment	Judgment
Numeric value of estimates not reported (p value only, significance or non-significance only)		Minor flaw-moderate risk
Mean only reported without p value or variance		Poor reporting
Estimate value reported with measure of spread around it and/or actual significance values		Low risk of bias

Sample size justification

Responses	Quote/ comment	Judgment
Not reported		Poor reporting
Justification by authors is incomplete or inaccurate		Minor flaw-moderate risk
Post-hoc analyses to justify sample size		Minor flaw-moderate risk
Well justified a priori		Low risk of bias

Appendix C: table of excluded reviews and studies, references awaiting assessment

Table of excluded reviews

Review ID	Reason
Flynn, 2012	Exposure only Chromium
Sinczuk-Walczak, 1996	FT assessed by native polish speaking neuroscientist. Not an SR a traditional review
Sjogren 1990	FT in Swedish. Recommendation document. No search methods

	or source
Sjogren, 1994	Multiple exposures of welding together; not Mn
Flynn, 2009	No search method or source
Guilarte, 2010	No search method or source
Jankovic, 2005	No search method or source
Lucchini 2000	No search method or source
Mergler 1997	No search method or source
Migliaccio 2004	No search method or source
Perl 2007	No search method or source
Santamaria, 2007	No search method or source
Winker, 2006	No search method or source
Zeidler-Erdely, 2012	No search method or source
Baranski, 1993	No search method or source
Bazylewicz-Walczak, 1996	No search method or source
Elbaz 2008	No search method or source
Finley, 2005	No search method or source
Rohling, 2007	No search method or source - meta-analysis of bowler studies as a critique
Wirth 2010	No search method or source, traditional review
Mathur, 2010	No search method or source; question is metals for male infertility; only 3 mouse studies for Mn
Valdes Hernandez Mdel 2012	Not about causality or association
Sjogren, 2004	Response to Antonini 2003
Antonini, 2003	Welding exposure only, no Mn
Szram, 2013	Welding exposure only, no Mn
Ambroise, 2006	Welding exposure only, no Mn
Moulin, 1997	Welding exposure, no FT but seems focused on hexa chromium, nickel and asbestos

Table of excluded primary studies

Study ID	Full Reference	Reason
Checkoway 2009	Checkoway H. Documenting neurotoxicity from occupational manganese exposure. Occupational and environmental medicine 2010;67(6):362-3 doi: 10.1136/oem.2009.047803[published Online First: Epub Date]|.	commentary
Albini 2007	Albini E, Benedetti L, Caruso A, et al. [Occupational exposure to manganese in ferroalloy industry: neurobehavioral effects in a workers' cohort]. Giornale italiano di medicina del lavoro ed ergonomia 2007;29(3 Suppl):272-4	cross sectional (no change from baseline)
Bowler 2003	Bowler RM, Gysens S, Diamond E, et al. Neuropsychological sequelae of exposure to welding fumes in a group of occupationally exposed men. International journal of hygiene and environmental health 2003;206(6):517-29	cross sectional (no change from baseline)
Bowler 2005	Bowler RM, Gysens S, Diamond E, et al. Manganese exposure: neuropsychological and neurological symptoms and effects in welders. Neurotoxicology 2006;27(3):315-26 doi: 10.1016/j.neuro.2005.10.007[published Online First: Epub Date]|.	cross sectional (no change from baseline)
Bowler 2006	Bowler RM, Roels HA, Nakagawa S, et al. Dose-effect relationships between manganese exposure and neurological, neuropsychological and pulmonary function in confined space bridge welders. Occupational and Environmental Medicine 2007;64(3):167-77 doi:	cross sectional (no change from baseline)

	10.1136/oem.2006.028761[published Online First: Epub Date]|.	
Chia 1993	Chia SE, Foo SC, Gan SL, et al. Neurobehavioural functions among workers exposed to manganese ore. SCANDINAVIAN JOURNAL OF WORK, ENVIRONMENT AND HEALTH 1993	cross sectional (no change from baseline)
Chia 1993	Chia SE, Goh J, Lee G, et al. Use of a computerized postural sway measurement system for assessing workers exposed to manganese. Clinical and experimental pharmacology & physiology 1993;20(9):549-53	cross sectional (no change from baseline)
Deschamps 2001	Deschamps FJ, Guillaumot M, Raux S. Neurological effects or workers exposed to manganese. JOURNAL OF OCCUPATIONAL AND ENVIRONMENTAL MEDICINE 2001	cross sectional (no change from baseline)
Dietz 2003	Dietz MC, Ihrig A, Bader M, et al. Occupational-medical field study of the chronic neurotoxicity of manganese dioxide. Arbeitsmedizin Sozialmedizin Umweltmedizin 2003;38(2):57-66	cross sectional (no change from baseline)
Dietz 2001	Dietz MC, Ihrig A, Wrazidlo W, et al. Results of magnetic resonance imaging in long-term manganese dioxide-exposed workers. Environmental research 2001;85(1):37-40 doi: 10.1006/enrs.2000.4068[published Online First: Epub Date]|.	cross sectional (no change from baseline)
Gennart 1992	Gennart JP, Buchet JP, Roels H. Fertility of male workers exposed to cadmium, lead or manganese. American journal of epidemiology 1992	cross sectional (no change from baseline)
Hochberg 1996	Hochberg F, Miller G, Valenzuela R, et al. Late motor deficits of Chilean manganese miners: a blinded control study. Neurology 1996;47(3):788-95	cross sectional (no change from baseline)
Iregren 1990	Iregren A. Psychological test performance in foundry workers exposed to low levels of manganese. Neurotoxicology and teratology 1990;12(6):673-5	cross sectional (no change from baseline)
Johnsen 2008	Johnsen HL, Soyseth V, Hetland SM, et al. Production of silicon alloys is associated with respiratory symptoms among employees in Norwegian smelters. International archives of occupational and environmental health 2008;81(4):451-59	cross sectional (no change from baseline)
Laohaudomchok 2010	Laohaudomchok W, Lin X, Herrick RF, et al. Neuropsychological effects of low-level manganese exposure in welders. Neurotoxicology 2011;32(2):171-9 doi: 10.1016/j.neuro.2010.12.014[published Online First: Epub Date]|.	cross sectional (no change from baseline)
Lucchini 1995	Lucchini R, Selis L, Folli D, et al. Neurobehavioral effects of manganese in workers from a ferroalloy plant after temporary cessation of exposure. Scandinavian journal of work, environment & health 1995;21(2):143-49	cross sectional (no change from baseline)
Park 2009	Park RM, Bowler RM, Roels HA. Exposure-response relationship and risk assessment for cognitive deficits in early welding-induced manganism. Journal of occupational and environmental medicine / American College of Occupational and Environmental Medicine 2009;51(10):1125-36 doi: 10.1097/JOM.0b013e3181bd8114[published Online First: Epub Date]|.	cross sectional (no change from baseline)
Recette 2012	Racette BA, Criswell SR, Lundin JI, et al. Increased risk of parkinsonism associated with welding exposure. Neurotoxicology 2012;33(5):1356-61 doi: 10.1016/j.neuro.2012.08.011[published Online First: Epub Date]|.	cross sectional (no change from baseline)
Rodier 1955	Rodier J. Manganese Poisoning in Moroccan Mines. British Journal of Industrial Medicine, Vol 12, No 1, pages 21-35, 17 references; 1955	cross sectional (no change from baseline)
Summers 2011	Summers MJ, Summers JJ, White TF, et al. The effect of occupational exposure to manganese dust and fume on neuropsychological functioning in Australian smelter workers. Journal of clinical and experimental neuropsychology 2011;33(6):692-703 doi: 10.1080/13803395.2011.553585[published Online First: Epub Date]|.	cross sectional (no change from baseline)
Wang 1989	Wang JD, Huang CC, Hwang YH, et al. Manganese induced parkinsonism: an outbreak due to an unrepaired ventilation control system in a ferromanganese smelter. British journal of industrial medicine 1989;46(12):856-9	cross sectional (no change from baseline)
Wennberg 1991	Wennberg A, Iregren A, Struwe G, et al. Manganese exposure in steel	cross sectional (no change from

	smelters a health hazard to the nervous system. Scandinavian journal of work, environment & health 1991;17(4):255-62	baseline)
Yuan 2006	Yuan H, He S, He M, et al. A comprehensive study on neurobehavior, neurotransmitters and lymphocyte subsets alteration of Chinese manganese welding workers. Life sciences 2006;78(12):1324-8 doi: 10.1016/j.lfs.2005.07.008[published Online First: Epub Date]|.	cross sectional (no change from baseline)
Zou 2014	Zou Y, ZouQing L, Zeng X, et al. Cognitive function and plasma BDNF levels among manganese-exposed smelters. Occupational and environmental medicine 2014;71(3):189-94	cross sectional (no change from baseline)
Beikuefner 1959	Beikuefner HD, Langhof H. [On the accumulated pathological ejaculate findings in electro-welders]. Das Deutsche Gesundheitswesen 1959;14:2280-8	cross sectional; no control
Thompson 2006	Thompson ML, Myers JE. Evaluating and interpreting exposure-response relationships for manganese and neurobehavioral outcomes. Neurotoxicology 2006;27(2):147-52 doi: 10.1016/j.neuro.2005.08.001[published Online First: Epub Date]|.	Modeling study based on Myers 2003 data
Zheng 2002	Zheng YX, Chan P, Pan ZF, et al. Polymorphism of metabolic genes and susceptibility to occupational chronic manganism. Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals 2002;7(4):337-46 doi: 10.1080/13547500210146740[published Online First: Epub Date]|.	No Mn measured
Dick 2006	Dick FD, De Palma G, Ahmadi A, et al. Environmental risk factors for Parkinson's disease and parkinsonism: the Geoparkinson study. Occupational and environmental medicine 2007;64(10):666-72 doi: 10.1136/oem.2006.027003[published Online First: Epub Date]|.	No Mn measured
El-Zein 2003	El-Zein M, Malo JL, Infante-Rivard C, et al. Incidence of probable occupational asthma and changes in airway calibre and responsiveness in apprentice welders. The European respiratory journal 2003;22(3):513-8	No Mn measured
Fryzek 2005	Fryzek JP, Hansen J, Cohen S, et al. A cohort study of Parkinson's disease and other neurodegenerative disorders in Danish welders. Journal of occupational and environmental medicine / American College of Occupational and Environmental Medicine 2005;47(5):466-72	No Mn measured
Gorell 1997	Gorell JM, Johnson CC, Rybicki BA, et al. Occupational exposures to metals as risk factors for Parkinson's disease. Neurology 1997;48(3):650-8	No Mn measured
Gorell 2004	Gorell JM, Peterson EL, Rybicki BA, et al. Multiple risk factors for Parkinson's disease. Journal of the neurological sciences 2004;217(2):169-74	No Mn measured
Gresham 1986	Gresham LS, Molgaard CA, Golbeck AL, et al. Amyotrophic Lateral Sclerosis and Occupational Heavy Metal Exposure: A Case Control Study. Neuroepidemiology, Vol 5, No 1, pages 29-38, 28 references; 1986	No Mn measured
Gustavsson 2000	Gustavsson P, Jakobsson R, Nyberg F, et al. Occupational exposure and lung cancer risk: a population-based case-referent study in Sweden. American journal of epidemiology 2000;152(1):32-40	No Mn measured
Marsh 2006	Marsh GM, Gula MJ. Employment as a welder and Parkinson disease among heavy equipment manufacturing workers. Journal of occupational and environmental medicine / American College of Occupational and Environmental Medicine 2006;48(10):1031-46 doi: 10.1097/01.jom.0000232547.74802.d8[published Online First: Epub Date]|.	No Mn measured
Park 2005	Park J, Yoo CI, Sim CS, et al. Occupations and Parkinson's disease: a multi-center case-control study in South Korea. Neurotoxicology 2005;26(1):99-105 doi: 10.1016/j.neuro.2004.07.001[published Online First: Epub Date]|.	No Mn measured
Park 2004	Park J, Yoo CI, Sim CS, et al. Occupations and Parkinson's disease: a case-control study in South Korea. Industrial health 2004;42(3):352-8	No Mn measured
Spinelli 1997	Spinelli A, Figa-Talamanca I, Osborn J. Time to pregnancy and occupation in a group of Italian women. International journal of	No Mn measured

	epidemiology 1997;26(3):601-9	
Tse 2012	Tse LA, Yu IT, Qiu H, et al. Occupational risks and lung cancer burden for Chinese men: a population-based case-referent study. Cancer causes & control : CCC 2012;23(1):121-31 doi: 10.1007/s10552-011-9861-1[published Online First: Epub Date]|.	No Mn measured
Nicolle-Mir 2010	Nicolle-Mir L. Occupational factors and Parkinson's disease. Environnement, Risques et Sante 2010;9(5):382-83	No Mn measured
Lundin 2011	Lundin J, Checkoway H, Criswell S, et al. Increased risk of parkinsonism associated with cumulative hours of welding. Occupational and Environmental Medicine 2011;68:A7	No Mn measured but authors say they are quantifying Mn in the ongoing study
Baldwin 2008	Baldwin M, Bouchard M, Larribe F, et al. Past occupational exposure to airborne manganese in a manganese alloy plant. Journal of occupational and environmental hygiene 2008;5(7):426-37 doi: 10.1080/15459620802115831[published Online First: Epub Date]|.	no outcome assessment
Feldman 1992	Feldman RG. Manganese as possible ecoetiologic factor in Parkinson's disease. Annals of the New York Academy of Sciences 1992;648:266-7	Not an empirical study
Crump 1999	Crump KS, Rousseau P. Results from eleven years of neurological health surveillance at a manganese oxide and salt producing plant. Neurotoxicology 1999;20(2-3):273-86	A repeat crosssectional survey design because, of the original cohort for which Reols 1987 measured baseline values only about one 3rd is in the current cohort and separate analyses for the original cohorts is not available.
Wastensson 2012	Wastensson G, Sallsten G, Bast-Pettersen R, et al. Neuromotor function in ship welders after cessation of manganese exposure. International archives of occupational and environmental health 2012;85(6):703-13 doi: 10.1007/s00420-011-0716-6[published Online First: Epub Date]|.	No Mn. Exposure is magnetic moment. Which is the correlation between iron and manganese multiplied to number of years at job. Because no real manganese values in air were available.

References awaiting full text assessment

1. Ahn SS, Lee KM. NEUROTOXICITY OF CHRONIC MANGANESE EXPOSURE CAUSING FRONTAL LOBE DYSFUNCTION. 1998
2. Anonymous. HEALTH EFFECTS OF MANGANESE EXPOSURE FOUND FROM WELDING IN CONFINED SPACE. OCCUPATIONAL SAFETY AND HEALTH REPORTER 2007
3. Backstrom I, Fryk G, Jakobsson R, et al. [Impaired pulmonary function after welding of painted steel]. Lakartidningen 1990;87(44):3622-3
4. Bleich S, Degner D, Sprung R, et al. Chronic manganism: fourteen years of follow-up. The Journal of neuropsychiatry and clinical neurosciences 1999;11(1):117
5. Cascini F, Daniele E, Gazzerro F. [Morbidity of electric welders in 2 industrial complexes]. Folia medica Folia medica (Naples, Italy) 1969;52(5):282-96
6. Cassan G, Kofman J, Gilly J, et al. [Pulmonary siderosis in electric welders: diffuse interstitial fibrosis and conglomerative form]. Archives des maladies professionnelles de medecine du travail et de securite sociale 1972;33(4):193-5
7. Checkoway H. Project 2: Parkinsonism and Metal Exposures Among Welders.
8. Flinn RH, Neal PA, Reinhart WH, et al. Chronic Manganese Poisining in an Ore-Crushing Mill. Division of Industrial Hygiene, National Institute of Health, US Public Health Service, Washington, DC, Public Health Bulletin, No 247, 77 pages, 34 references; 1940 1940
9. Gibbs JP, Crump KS, Houck DP, et al. Focused medical surveillance: a search for subclinical movement disorders in a cohort of U.S. workers exposed to low levels of manganese dust. Neurotoxicology 1999;20(2-3):299-313
10. Gorell JM, Johnson CC, Rybicki BA, et al. Occupational exposure to manganese, copper, lead, iron, mercury and zinc and the risk of Parkinson's disease. Neurotoxicology 1999;20(2-3):239-47
11. He SC, Niu Q. Subclinical neurophysiological effects of manganese in welding workers. International journal of immunopathology and pharmacology 2004;17(2 Suppl):11-6
12. Horvitz I, Uiberall E. Neuropsychiatric Considerations of Managanese Intoxication. Revista Neurologica de Buenos Aires, Vol IX, No 1, 48 pages, 31 references; 1944 1944
13. Jonderko G, Kujawska A, Langauer-Lewowicka H. Problems of chronic manganese poisoning on the basis of investigations of workers at a manganese alloy foundry. Internationales Archiv fur Arbeitsmedizin 1971;28(3):250-64
14. Lu YR, Li YN, Li HJ. [Investigation on occupational manganese poisoning hazard in Guang Xi]. Zhonghua lao dong wei sheng zhi ye bing za zhi = Zhonghua laodong weisheng zhiyebing zazhi = Chinese journal of industrial hygiene and occupational diseases 2009;27(10):616-7
15. Lucchini R, Apostoli P, Perrone C, et al. Long-term exposure to "low levels" of manganese oxides and neurofunctional changes in ferroalloy workers. Neurotoxicology 1999;20(2-3):287-97
16. Mikhailova A, Zaprianov Z, Antov G, et al. [The assessment of the occupational risk to workers in electric steel production]. Problemi na khigienata 1995;20:122-8
17. Mir L. Dose-effect relationship between manganese exposure and neurological, neuropsychological and pulmonary function in confined space bridge welders. Environnement, Risques et Sante 2007;6(5):333-34
18. Mora LB. Long-term change in attention, working memory, anxiety and depression due to confined space welding and manganese exposure, 2012.
19. Nikolov Z. [Hearing reduction caused by manganese and noise]. JFORL Journal francais d'oto-rhino-laryngologie; audiophonologie et chirurgie maxillo-faciale 1974;23(3):231-4

20. Niu Q, Shuchang H, Sheng W, et al. Neurobehavioral functions, serum prolactin and plasma renin activity of manganese-exposed workers. International journal of immunopathology and pharmacology 2004;17(2 Suppl):17-24
21. Petkova V, Karadjov K. ON THE CLINICAL MANIFESTATIONS AND DIAGNOSIS OF MANGANESE PARKINSONISM. GIGIENA TRUDA I PROFESSIONALNYE ZABOLEVANIYA
22. Petkova V, Karadzhov K. [The clinical picture and diagnosis of manganese-induced parkinsonism]. Gigiena truda i professional'nye zabolevaniia 1991(4):16-8
23. Saakadze VP, Vasilov BG. Occupational bronchial asthma due to manganese. Gigiena truda i professional'nye zabolevanija; Moskva, USSR, June 1977, No6, p19-23 7 ref 1977
24. Saric M, Lucic-Palaic S, Paukovic R, et al. The effect of manganese on the respiratory system. Arhiv za higijenu rada i toksikologiju; Zagreb, Yugoslavia, 1974, Vol25, No1, p15-26 Illus 9 ref 1974
25. Steenland K, Beaumont J, Hornung R. The use of regression analyses in a cohort mortality study of welders. Journal of chronic diseases 1986;39(4):287-94
26. van Beukering JA. [The incidence of pneumonia among miners in an iron and a manganese mine in South Africa]. Nederlands tijdschrift voor geneeskunde 1966;110(10):473-4
27. Yoshida S. [Pneumoconiosis in manganese refinery workers]. Nihon Igaku Hoshasen Gakkai zasshi Nippon acta radiologica 1987;47(8):1027-33
28. Zheng Y, He F, Chan P, et al. [Genetic polymorphism and susceptibility to occupational chronic manganism: a case-control study]. Zhonghua yu fang yi xue za zhi [Chinese journal of preventive medicine] 1999;33(2):78-80
29. Zheng Y, Zhou X, Pan J, et al. [Changes of neurobehavioral functions in workers exposed to manganese]. Wei sheng yan jiu = Journal of hygiene research 1999;28(4):198-9, 202
30. Zschiesche W. Longitudinal occupational health study of special steel production workers, 1996.

Appendix D: table of characteristics of included reviews

Table of characteristics of included reviews

Study ID	Aim/ question	Population	Exposure	Comparison	Outcome	Study designs included	Search & selection	No. of studies included	No. Included in Meta-analysis (MA)
Assem, 2011	An up-to-date analysis of the available published studies on the carcinogenic and genotoxic potential of inorganic Mn	Any : cells, bacteria, animal, human	Welding fumes containing inorganic Mn	NR	Cancer, mortality, measures of DNA damage	NR, likely any	Criteria document (IEH 2004) and a systematic search- no description of search or selection methods.	NR	MA not done
Bailey, 2009	Provide NOAEL values and revised Uncertainty Factors for calculating an Mn reference concentration	Workers	Mn dust (not welding fumes) in personal air	exposed vs. unexposed	Neurological effects	Comparative studies: exp vs non exp	PubMed	12 studies of 8 cohorts	MA not done
Baker 2014	(1) to review : blood as a biomarker for Mn exposure; quantitative relationships between MnA and MnB reported in the literature	NR, likely any	Mn exposure, including serum, plasma, urine, and saliva, and brain tissue, toenails, and hair.	various levels of exposure: MnA vs MnB	Mean, SD	NR, likely any	PubMed: search terms given. English only	29	26 in regression analysis
Greiffenstein, 2007	To examine the association between specific neurobehavioral measures and various variable classes such as demographics	Workers	Occupational mn exposure	NR	Mean, SD	Comparative studies	Lees-Haley et al. (2006). PsycInfo, PubMed: search terms given.	19	19
Hobson, 2011	To develop and validate a multivariate model to estimate quantitative levels of welding fume exposures based on welding particulate mass and Mn concentrations reported in the published literature	Welders	Mn personal air exposure in welding	various levels of exposure	Mean welding fume levels, minimum 6 hrs long	Any descriptive	Medline: Search terms given, English only	27	27
Lees-Haley,	To perform a meta-analysis of the quantitative	Workers, occupationall	Current or past chronic Mn	NR	Common neuropsychologic	Comparative studies	PubMed, Medline, PsycInfo, references of papers and	25	25

2006	empirical literature to determine the effects of occupational exposure to manganese on neuropsychological functioning	y exposed	exposure		al measures of cognitive, sensory, motor and psychological function		book chapters		
Li, 2014	To evaluate the sensitivity, feasibility, and effectiveness of the Pallidal Index (PI) as a biomarker of brain manganese(Mn) accumulation	Workers	Occupational exposure to Mn for a work period of 8h/d, 5d/wk.	Mn exposed vs. healthy subjects without exposure to Mn concentrations exceeding 0.15mg/m 3 of the time-weighted average	Pallidal index	Retrospective study, clinical trial, quasi-randomized controlled trial	MEDLINE, Cochrane Library, Embase, Chinese Biological Medical Literature (CBM), Chinese National Knowledge Infrastructure (CNKI),Chinese Wanfang, and Chongqing VIP databases for epidemiological studies. Reference of included and author contact for unpublished studies. Independent and duplicate DE with consensus of the two as final.	10 (16 reports/studies of the same 10 cohorts/subsamples)	8
Ma, 2011	To evaluate the effects of occupational manganese exposure on neurobehavioral function	Workers exposed to manganese	Occupational Mn: air, urine, or blood	unclear	WHO.NCTB or NES2C3 or Other methods to assess neurobehavioral function.	Randomized controlled study type	Qinghua Fang, Chongqing VIP, PubMed and other databases, 1990 to 2010.	24	none
Meyer-Baron, 2009	1. Is there consistent evidence for a negative impact of occupational exposure to Mn on performance? 2. Which functions are affected? 3. Which performance tests are sensitive to mirror the impact of Mn exposure? 4. Are the effects related to indices of exposure?	NR	Occupational exposure to Mn as Mn blood or Mn inhaled	exposed vs. 'control' probably means unexposed	Standardized neuropsychological tests of cognitive and motor function used in more than one study	Epidemiological study	PubMed, science direct, web of science	13	at least 3 and at most 9 in an MA
Meyer-Baron, 2013	IPD analysis of 2009 MA to find: 1) neuro effects of Mn when considering confounding. 2) dose response from Mn blood	Same as 2009	Occupational Mn exposure	control' probably means unexposed	Standardized neuropsychological tests of cognitive and motor function	Comparative studies	PubMed, Medline, PsycInfo, references of papers and book chapters. Limited to 2009 only	8	8

	marker 3) individual susceptibility factors for Mn effects				which must be used in more than one study				
Mortimer, 2012	To examine associations of welding and manganese exposure with Parkinson disease (PD) using meta-analyses of data from cohort, case-control, and mortality studies	Welders, cutters	NR, tables indicate welding or Mn	NR	Clinically diagnosed Parkinson's disease	Cohort, case-control, and mortality studies	PubMed, CDSR and other published reviews	13	11 (9 in one and 3 in the other MA)
Sutedja, 2009	To evaluate existing evidence on whether life time exposure to chemicals or heavy metals increase the risk of developing ALS (amyotrophic lateral sclerosis)	Any persons	A chemical agent or a metal including Mn	NR, likely exposed vs. unexposed	Sporadic ALS (not endemic)	Case-control or cohort	Medline, CINAHL, EMBASE, Cochrane databases. Related article searches in PubMed and Web of science	relevant to Mn 1, total 10	MA not done
Zoni, 2007	The aim of this study was to update and compare the different neurobehavioral tests used in the Mn literature, in order to identify the most sensitive methods to be included in future research studies.	Any adults or children	NR, likely any Mn exposure (occupational, non-occupational)	one test of neurobehavioral versus another	Tests of pre-clinical neurobehavioral effects of Mn	Any	Published from 1986 Medline using PubMed.	31 total, 18 occupational, 7 adults in environment and 6 in children	MA not done
IEA 2004	Review the potential health effects of inorganic manganese compounds; it proposes a new health-based Occupational Exposure Limit (OEL) for Mn	Workers	Any occup exposure to Mn	Any comparison	All health outcomes	Any	Medline, Embase, Toxfile, Datastar, NIOSHTIC include databases	28 for neurological tests	MA not done

Table of characteristics of included reviews (continued)

Study ID	Primary effect measure	Method of synthesis	Heterogeneity exploration	Publication bias assessment	Results obtained	Authors conclusions	Conflict of interest/ sponsor
Assem, 2011	NR	NR, likely vote counting	NR	NR	NR	Insufficient evidence to indicate that inorganic Mn exposure produces cancer in animals or humans	Authors thank the Manganese Health Research Program (MHRP) for funding the original criteria document on which their paper is based
Bailey, 2009	NR	NR	NR	NR	NOAEL=2 microgramsper meter cube per week; BDML= 7 micrograms per meter cube per week	Exposure limit should be inflated to 2-7 microg/m3 of inhaled Mn at work(per week)	One of the authors (Beck) has been named as an expert in litigation involving air exposures to manganese, among other constituents. Some of the underlying work for this manuscript was conducted in the context of an assignment from an industrial client. Preparation of the manuscript was not supported by any client and the opinions are solely those of the authors.
Baker 2014	correlation coefficient R square, slope measurement	segmented regression, non-linear least-squares estimation	outliers removed from analysis	NR	The R2 value for the segmented regression= 0.34 considering all points. Simple weighted linear regression model yields a lower R2= 0.25	There is a point above which Mn levels in the blood begin to act as an exposure biomarker for inhaled Mn. The primary effect was present at higher levels of exposure, such as those seen among ferroalloy and Mn smelting operations, and production of Mn-containing minerals. Welders typically have intermediate levels of Mn exposures, in the range of 10 – 200 µg/m3.	NR
Greiffenstein, 2007	Cohen's d (SMD) for continuous and OR for categorical data	fixed effect MA	probably a meta-regression was done. Sensitivity analysis by excluding the one where education was force matched	NR	SMD (95%CI) for clerical substitution test: −0.52 (−0.66 to −0.39); SMD (95%) For digit tapping: −0.46 (−0.59 to−0.32). [SI=Nearly half of the tests (outcomes) show significant negative effect]	The data did not support a theory of preclinical (“early”) neuromotor or cognitive dysfunction. Overall, the pooled data are more consistent with covariate effect than toxic effect, insofar as the pooled exposure group showed demographics less favorable to neuropsychological performance than the pooled	Dr. Greiffenstein was a paid consultant to the manganese consumable industry in the past but not during any portion of this study. Dr. Lees-Haley is presently consulting to attorneys representing current and former manufacturers of welding consumables.

						referent groups. Future consideration of demographic and biological covariates is necessary before inferring subtle toxin-induced brain damage because neuropsychological tests are nonspecific.	
Hobson, 2011	R-square, mean predicted exposure for welding particulate and Mn component within it.	weighted regression: natural log of Mn level as the dependent variable in the model	ventilation level explored for effect size variation	NR	R square=3.23; Mn makes 4% (95%CI=2.1 to 7.2% of total welding fumes in welders.	The model is useful in the absence of individually measured historically collected data	Funded by Michael Fox foundation, national institutes of health grants (r01 es013743, k24 es017765, p42 es04696); the clinical science translational award (ncrr ul1 rr024992); the neuroscience blueprint grant (ns057105); the American Parkinson disease association advanced research center at Washington university; the greater St Louis chapter of the APDA.
Lees-Haley, 2006	SMD (Cohen's d)	unweighted average and weighted meta-analysis	Q statistic. Meta-regression for definitions of exposure	NR	Weighted MD=-0.17, SE=0.04	A small negative effect of Mn exposure exists, but it should be unmeasurable at individual level	First author consulted for lawyers of the welding consumables manufacturers
Li, 2014	WMD	random effects model when I square over 50% otherwise fixed	I square, meta-regression (NR on what variables)	funnel plot, eggers test	Effect of Pallidal index for brain accumulation - WMD= 7.76 (95%CI, 4.86 to 10.65) I2: 85.7%. Egger's test p=0.014	PI considered as a sensitive, feasible, effective and semi-quantitative index in evaluating brain Mn accumulation. However, the results should be interpreted with caution.	Declared that no competing interests exist.
Ma, 2011	Standardized mean difference	Planned MA in RevMan and regression in SPSS indicated in methods but only narrative synthesis in results	For obvious heterogeneity (P <0.05) in 7 tests, selected random effects model. Possibly boot strapping.	NR, although paper mentions it may have publication bias	The effect size range is -0. 10 to 1.48. Effect size was negatively correlated with other indicators; Rz values were 0.55, 0.09, 0.32, which Simple reaction time test indicator P value of 0.09. P values were greater for the other indexes. Cumulative exposure levels and neurobehavioral function effects values are also negatively correlated.	This suggests that cumulative exposure of workers to some extent reflects the level of long-term exposure of workers, but still cannot accurately represent the manganese in the central nervous system. The dose-effect relationship is not obvious.	None to be declared
Meyer-Baron, 2009	(d)weighted mean of all effects	random effects MA	between study variance = 93 to 97%; meta	NR	For six outcomes the range was significant between d= -0.23 to -0.36	There is an overall negative effect of Mn exposure in air on motor functions; blood Mn not useful as	None to be declared

			regression on exposure measures- MnInh and MnB			biomarker	
Meyer-Baron, 2013	MD for Simple reaction time(SRT)	fixed effect ANOVA	ANCOVA on various factors e.g. age, education,	NR	MD SRT=7.8;p<0.01, 9 of the 30 tests significant for Mn exposed versus control participants	Confirmed 2009 findings. Lower cognitive and motor performance seen in a heterogeneous sample of Mn exposed workers	None to be declared
Mortimer, 2012	RR	both fixed and random effects MA	I square statistic, but heterogeneity not explored	NR	RR=0.86, 95%CI= 0.8 to 0.92	Welding and Mn exposure are not associated with Parkinson's disease	First author and the early literature search for this paper was paid by welding industry defense group
Sutedja, 2009	OR,SMR,SIR or PMR	narrative with tabulation and graphical presentation of effect sizes and 95%CI	because of heterogeneity the MA was not performed	NR	No significant association for Mn in figure, values NR	This SR had difficulty in attaining a high level of evidence due to lack of high quality of methodological and exposure components in studies.	Authors declare there are no conflicts of interest.
Zoni, 2007	NR	NR, likely vote counting	NR	NR	Good tests for future studies are: Intellectual abilities: Standard progressive matrices or WAIS-R; Mood & symptoms: POMS or mood scale BSI; Motor& speed: Finger tapping, visual reaction time, pursuit aiming, pegboard Test; Cognitive functions: Trail making test, symbol digit, digit span, addition test,Rey-15 item, or Benton visual retention or WMS Neurological functions: Tremor test, Luria-Nebraska motor battery, CATSYS system	Literature on manganese neurobehavioral effect is quite consistent; however, further improvement may be achieved by using better-structured and more comparable evaluation methods. A test battery is suggested.	NR, likely none: sponsor is the European Union; quote "This study reflects only the authors' views. The EU Community is not liable for any use that may be made of the information contained therein."
IEA 2004	NR	narrative	NR	NR	Exposure levels vary across and within the same work-sites; biomarkers are poorly correlated to Mn exposure; respiratory or cardiovascular effects cannot be ascribed accurately to Mn exposure; neurological motor deficit of sub clinical level can occur at low exposures.	NOAEL 1 mg/m3; limiting exposure to 0.1 mg/m3 respirable manganese will prevent the subtlest detectable effect. Pulmonary or cardiovascular effects do not appear to occur at levels below those at which identifiable neurological changes can be detected.	None reported

References to the literature

1. Food and Nutrition Board. Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc Washington, DC: Institute of Medicine, 2002.

2. Boudia N, Halley R, Kennedy G, et al. Manganese concentrations in the air of the Montreal (Canada) subway in relation to surface automobile traffic density. Sci Total Environ 2006;**366**(1):143-7 doi: 10.1016/j.scitotenv.2005.09.094[published Online First: Epub Date]|.

3. Fell JM, Reynolds AP, Meadows N, et al. Manganese toxicity in children receiving long-term parenteral nutrition. Lancet 1996;**347**(9010):1218-21

4. Department of Health and Human Services. NIOSH POCKET GUIDE TO CHEMICAL HAZARDS: DHHS (NIOSH) CDC 2007.

5. HaMai D, Bondy SC. Oxidative basis of manganese neurotoxicity. Ann N Y Acad Sci 2004;**1012**:129-41

6. Zhang S, Zhou Z, Fu J. Effect of manganese chloride exposure on liver and brain mitochondria function in rats. Environ Res 2003;**93**(2):149-57

7. Gavin CE, Gunter KK, Gunter TE. Manganese and calcium efflux kinetics in brain mitochondria. Relevance to manganese toxicity. Biochem J 1990;**266**(2):329-34

8. Bowman AB, Kwakye GF, Herrero Hernandez E, et al. Role of manganese in neurodegenerative diseases. J Trace Elem Med Biol 2011;**25**(4):191-203 doi: 10.1016/j.jtemb.2011.08.144[published Online First: Epub Date]|.

9. Racette BA, Aschner M, Guilarte TR, et al. Pathophysiology of manganese-associated neurotoxicity. Neurotoxicology 2012;**33**(4):881-6 doi: 10.1016/j.neuro.2011.12.010[published Online First: Epub Date]|.

10. Huang WH, Lin JL. Acute renal failure following ingestion of manganese-containing fertilizer. Journal of toxicology Clinical toxicology 2004;**42**(3):305-7

11. Sanchez B, Casalots-Casado J, Quintana S, et al. Fatal manganese intoxication due to an error in the elaboration of Epsom salts for a liver cleansing diet. Forensic science international 2012;**223**(1-3):e1-4 doi: 10.1016/j.forsciint.2012.07.010[published Online First: Epub Date]|.

12. Deng Q, Liu J, Li Q, et al. Interaction of occupational manganese exposure and alcohol drinking aggravates the increase of liver enzyme concentrations from a cross-sectional study in China. Environmental Health 2013;**12**(1):30

13. Lebda MA, El-Neweshy MS, El-Sayed YS. Neurohepatic toxicity of subacute manganese chloride exposure and potential chemoprotective effects of lycopene. Neurotoxicology 2012;**33**(1):98-104 doi: http://dx.doi.org/10.1016/j.neuro.2011.12.008|.

14. Bowler RM, Gocheva V, Harris M, et al. Prospective study on neurotoxic effects in manganese-exposed bridge construction welders. Neurotoxicology 2011;**32**(5):596-605 doi: 10.1016/j.neuro.2011.06.004[published Online First: Epub Date]|.

15. Wastensson G, Sallsten G, Bast-Pettersen R, et al. Neuromotor function in ship welders after cessation of manganese exposure. International archives of occupational and environmental health 2012;**85**(6):703-13 doi: 10.1007/s00420-011-0716-6[published Online First: Epub Date]|.

16. Bouchard M, Mergler D, Baldwin M, et al. Neurobehavioral functioning after cessation of manganese exposure: a follow-up after 14 years. American journal of industrial medicine 2007;**50**(11):831-40 doi: 10.1002/ajim.20407[published Online First: Epub Date]|.

17. Beuter A, Lambert G, MacGibbon B. Quantifying postural tremor in workers exposed to low levels of manganese. Journal of neuroscience methods 2004;**139**(2):247-55 doi: 10.1016/j.jneumeth.2004.05.002[published Online First: Epub Date]|.

18. Meyer-Baron M, Knapp G, Schaper M, et al. Performance alterations associated with occupational exposure to manganese--a meta-analysis. Neurotoxicology 2009;**30**(4):487-96 doi: 10.1016/j.neuro.2009.05.001[published Online First: Epub Date]|.

19. Hine CH, Pasi A. Manganese intoxication. The Western journal of medicine 1975;**123**(2):101-7

20. Saric M, Markicevic A, Hrustic O. Occupational exposure to manganese. British journal of industrial medicine 1977;**34**(2):114-8

21. Smyth LT, Ruhf RC, Whitman NE, et al. Clinical manganism and exposure to manganese in the production and processing of ferromanganese alloy. Journal of occupational medicine : official publication of the Industrial Medical Association 1973;**15**(2):101-9

22. Arias P, Robles-García V, Espinosa N, et al. Validity of the finger tapping test in Parkinson's disease, elderly and young healthy subjects: Is there a role for central fatigue? Clinical Neurophysiology 2012;**123**(10):2034-41 doi: http://dx.doi.org/10.1016/j.clinph.2012.04.001[published Online First: Epub Date]|.

23. Shamliyan T, Kane RL, Ansari MT, et al. Development of the Quality Criteria To Evaluate Nontherapeutic Studies of Incidence, Prevalence, or Risk Factors of Chronic Diseases: Pilot Study of New Checklists. . Methods Research Report AHRQ Publication No 11-EHC008-EF Agency for Healthcare Research and Quality, 2011.

24. Shamliyan TA, Kane RL, Ansari MT, et al. Development quality criteria to evaluate nontherapeutic studies of incidence, prevalence, or risk factors of chronic diseases: pilot study of new checklists. Journal of clinical epidemiology 2011;**64**(6):637-57 doi: http://dx.doi.org/10.1016/j.jclinepi.2010.08.006[published Online First: Epub Date]|.

25. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. Journal of clinical epidemiology 2009;**62**(10):1013-20 doi: 10.1016/j.jclinepi.2008.10.009[published Online First: Epub Date]|.

26. Shea BJ, Bouter LM, Peterson J, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). PLoS One 2007;**2**(12):e1350 doi: 10.1371/journal.pone.0001350[published Online First: Epub Date]|.

27. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;**7**:10 doi: 10.1186/1471-2288-7-10[published Online First: Epub Date]|.

28. Greenland S. Chapter 16 Applications of stratified analysis methods. In: Rothman K GS, Lash LT, ed. Modern epidemiology. third ed. Philadelphia, USA: Lippincott Williams and Wilkins, 2008.

29. Greenland S, Lash T. Chapter 19 Bias analysis. In: Rothman K GS, Lash LT, ed. Modern epidemiology. third ed. Philadelphia, USA: Lippincott Williams and Wilkins, 2008.

30. Assem FL, Holmes P, Levy LS. The mutagenicity and carcinogenicity of inorganic manganese compounds: a synthesis of the evidence. Journal of toxicology and environmental health Part B, Critical reviews 2011;**14**(8):537-70 doi: 10.1080/10937404.2011.615111[published Online First: Epub Date]|.

31. Bailey LA, Goodman JE, Beck BD. Proposal for a revised Reference Concentration (RfC) for manganese based on recent epidemiological studies. Regulatory toxicology and pharmacology : RTP 2009;**55**(3):330-9 doi: 10.1016/j.yrtph.2009.08.005[published Online First: Epub Date]|.

32. Baker MG, Simpson CD, Stover B, et al. Blood manganese as an exposure biomarker: state of the evidence. Journal of occupational and environmental hygiene 2014;**11**(4):210-7 doi: 10.1080/15459624.2013.852280[published Online First: Epub Date]|.

33. Greiffenstein MF, Lees-Haley PR. Neuropsychological correlates of manganese exposure: a meta-analysis. Journal of clinical and experimental neuropsychology 2007;**29**(2):113-26 doi: 10.1080/13803390600781105[published Online First: Epub Date]|.

34. Hobson A, Seixas N, Sterling D, et al. Estimation of particulate mass and MN exposure levels among welders. Occupational and Environmental Medicine 2011;**68**:A94-A95

35. Lees-Haley PR, Rohling ML, Langhinrichsen-Rohling J. A meta-analysis of the neuropsychological effects of occupational exposure to manganese. The Clinical neuropsychologist 2006;**20**(1):90-107 doi: 10.1080/13854040500203274[published Online First: Epub Date]|.

36. Li SJ, Jiang L, Fu X, et al. Pallidal index as biomarker of manganese brain accumulation and associated with manganese levels in blood: a meta-analysis. PLoS One 2014;**9**(4):e93900 doi: 10.1371/journal.pone.0093900[published Online First: Epub Date]|.

37. Ma XY, Zou YF, Yang XB. [A meta analysis on neuro-behavioral function of the workers exposed to manganism occupationally]. Zhonghua lao dong wei sheng zhi ye bing za zhi = Zhonghua laodong weisheng zhiyebing zazhi = Chinese journal of industrial hygiene and occupational diseases 2011;**29**(3):235-8

38. Meyer-Baron M, Schaper M, Knapp G, et al. Statistical means to enhance the comparability of data within a pooled analysis of individual data in neurobehavioral toxicology. Toxicology letters 2011;**206**(2):144-51 doi: 10.1016/j.toxlet.2011.07.003[published Online First: Epub Date]|.

39. Meyer-Baron M, Schaper M, Knapp G, et al. The neurobehavioral impact of manganese: results and challenges obtained by a meta-analysis of individual participant data. Neurotoxicology 2013;**36**:1-9 doi: 10.1016/j.neuro.2013.02.003[published Online First: Epub Date]|.

40. Mortimer JA, Borenstein AR, Nelson LM. Associations of welding and manganese exposure with Parkinson disease: review and meta-analysis. Neurology 2012;**79**(11):1174-80 doi: 10.1212/WNL.0b013e3182698ced[published Online First: Epub Date]|.

41. Sutedja NA, Veldink JH, Fischer K, et al. Exposure to chemicals and metals and risk of amyotrophic lateral sclerosis: A systematic review. Amyotrophic Lateral Sclerosis 2009;**10**(5-6):302-09

42. Zoni S, Albini E, Lucchini R. Neuropsychological testing for the assessment of manganese neurotoxicity: a review and a proposal. American journal of industrial medicine 2007;**50**(11):812-30 doi: 10.1002/ajim.20518[published Online First: Epub Date]|.

43. Institute for Environment and Health/Institute of Occupational Medicine. Occupational exposure limits: Criteria document for manganese and inorganic manganese compounds. IEH Web Report Leicester: MRC Institute for Environment and Health, 2004.

44. Blond M, Netterstrom B. Neuromotor function in a cohort of Danish steel workers. Neurotoxicology 2007;**28**(2):336-44 doi: 10.1016/j.neuro.2006.07.010[published Online First: Epub Date]|.

45. Blond M, Netterstrom B, Laursen P. Cognitive function in a cohort of Danish steel workers. Neurotoxicology 2007;**28**(2):328-35 doi: 10.1016/j.neuro.2007.01.006[published Online First: Epub Date]|.

46. Boojar MMA, Goodarzi F. A longitudinal follow-up of pulmonary function and respiratory symptoms in workers exposed to manganese. Journal of Occupational and Environmental Medicine 2002

47. Bouchard M, Mergler D, Baldwin M, et al. Neuropsychiatric symptoms and past manganese exposure in a ferro-alloy plant. Neurotoxicology 2007;**28**(2):290-7 doi: 10.1016/j.neuro.2006.08.002[published Online First: Epub Date]|.

48. Bouchard M, Mergler D, Baldwin ME, et al. Manganese cumulative exposure and symptoms: a follow-up study of alloy workers. Neurotoxicology 2008;**29**(4):577-83 doi: 10.1016/j.neuro.2008.04.013[published Online First: Epub Date]|.

49. Fored CM, Fryzek JP, Brandt L, et al. Parkinson's disease and other basal ganglia or movement disorders in a large nationwide cohort of Swedish welders. Occup Environ Med 2006;**63**(2):135-40 doi: 10.1136/oem.2005.022921[published Online First: Epub Date]|.

50. Hobbesland Å, Kjuus H, Thelle DS. Study of cancer incidence among 6363 male workers in four Norwegian ferromanganese and silicomanganese producing plants. OCCUPATIONAL AND ENVIRONMENTAL MEDICINE 1999

51. Mergler D, Huel G, Bowler R, et al. Nervous system dysfunction among workers with long-term exposure to manganese. Environ Res 1994;**64**(2):151-80 doi: 10.1006/enrs.1994.1013[published Online First: Epub Date]|.

52. Park J, Yoo CI, Sim CS, et al. A retrospective cohort study of Parkinson's disease in Korean shipbuilders. Neurotoxicology 2006;**27**(3):445-9 doi: 10.1016/j.neuro.2006.01.005[published Online First: Epub Date]|.

53. Park RM, Bouchard MF, Baldwin M, et al. Respiratory manganese particle size, time-course and neurobehavioral outcomes in workers at a manganese alloy production plant. Neurotoxicology 2014 doi: 10.1016/j.neuro.2014.03.015[published Online First: Epub Date]|.

54. Roels HA, Ortega Eslava MI, Ceulemans E, et al. Prospective study on the reversibility of neurobehavioral effects in workers exposed to manganese dioxide. Neurotoxicology 1999;**20**(2-3):255-71

55. Lundin J, Checkoway H, Criswell S, et al. Increased risk of parkinsonism associated with cumulative hours of welding. Occupational and Environmental Medicine 2011;**68**:A7

56. Bowler RM, Nakagawa S, Drezgic M, et al. Sequelae of fume exposure in confined space welding: a neurological and neuropsychological case series. Neurotoxicology 2007;**28**(2):298-311 doi: 10.1016/j.neuro.2006.11.001[published Online First: Epub Date]|.

57. Calne DB, Chu NS, Huang CC, et al. Manganism and idiopathic parkinsonism: similarities and differences. Neurology 1994;**44**(9):1583-6

58. Olanow CW. Manganese-induced parkinsonism and Parkinson's disease. Ann N Y Acad Sci 2004;**1012**:209-23

59. McMillan G. Is electric arc welding linked to manganism or Parkinson's disease? Toxicological reviews 2005;**24**(4):237-57

60. Myers JE, Fine J, Ormond-Brown D, et al. Estimating the prevalence of clinical manganism using a cascaded screening process in a South African manganese smelter. Neurotoxicology 2009;**30**(6):934-40 doi: 10.1016/j.neuro.2009.08.004[published Online First: Epub Date]|.

61. Camicioli R, Grossmann SJ, Spencer PS, et al. Discriminating mild parkinsonism: methods for epidemiological research. Movement disorders : official journal of the Movement Disorder Society 2001;**16**(1):33-40

62. Lezak MD. *Neuropsychological Assessment*: Oxford University Press, 2004.

63. Keski-Santti P, Kovala T, Holm A, et al. Quantitative EEG in occupational chronic solvent encephalopathy. Human & experimental toxicology 2008;**27**(4):315-20 doi: 10.1177/0960327107082231[published Online First: Epub Date]|.

64. Keski-Santti P, Holm A, Akila R, et al. P300 of auditory event related potentials in occupational chronic solvent encephalopathy. Neurotoxicology 2007;**28**(6):1230-6 doi: 10.1016/j.neuro.2007.08.004[published Online First: Epub Date]|.

65. Vandenplas O. Occupational asthma: etiologies and risk factors. Allergy, asthma & immunology research 2011;**3**(3):157-67 doi: 10.4168/aair.2011.3.3.157[published Online First: Epub Date]|.

66. Papapetropoulos S, Katzen HL, Scanlon BK, et al. Objective quantification of neuromotor symptoms in Parkinson's disease: implementation of a portable, computerized measurement tool. Parkinson's disease 2010;**2010**:760196 doi: 10.4061/2010/760196[published Online First: Epub Date]|.

67. van Tongeren M, Ritchie P, Galea K, et al. Development of guidance and tools to facilitate the use of routine exposure data in future epidemiology studies of manganese exposed workers: Institute of Occupational Medicine Research Park North Riccarton Edinburgh, UK 2009:23.

68. Maroni M. *Study of toxic effects on the central and peripheral nervous system of workers in the ferro-alloy industry*. Luxembourg: Office for Official Publications of the European Communities, 1997.

69. International Mangnese Institute. Manganese Production. Secondary Manganese Production 2007. http://www2.manganese.org/production.php.

Evaluation of the "Review of occupational exposure to manganese and the potential health effects of such exposure" by Ijaz et al.

Rotterdam, October 12, 2014
Prof Dr Alex Burdorf
Erasmus MC Rotterdam
The Netherlands

Overall assessment

The review presents a detailed, well-thought overview of the available information on health effects of occupational exposure to manganese. As such, this review clearly demonstrates the problems in appreciation of the potential risks of manganese exposure and the lack of clear evidence to arrive at solid conclusions.

A major point of concern is that the description of bias in studies does not link well with the theory of attenuation of exposure-response associations due to poor exposure characterization. This becomes already clear in differences in definition. In exposure assessment one would differentiate between random misclassification / measurement error and bias as systematic (non-random) difference. In the current review bias can have both a systematic and random component. It would help to clearly express that the appraisal of bias in this review is about both systematic bias as well as attenuation as bias mechanism. Hence, in several parts of the manuscript the consequences of measurement strategy for attenuation could be given more attention.

Response:

Thank you or your comments and for appreciation of our work. We are happy that the message on poor exposure assessment in current literature come across clearly.

We however do not think that measurement error and misclassification are always random. In our opinion both of these contribute to a systematic bias in a study which then biases the results of a review if not clearly identified and accounted for. We are not sure what attenuation means here: is it the attenuation of exposure which is considered always to be a result of measurement error and random misclassification? If so we do not think that is the case.

If however attenuation refers to that of our findings and conclusions which are a result of accounting for the various biases, we agree that it should be given more attention however that would lengthen the review discussion.

Specific remarks

1. Summary:
   - authors should consider carefully the terminology used, eg the available information from epidemiological studies will allow evaluation of the exposure-response relationship (i.e. exposure measurements in air) as well as dose-response relationship (i.e. measurements in human material such as urine). This important distinction is unclear in the summary statement. In fact, the term exposure and dose are not used consistently throughout the complete manuscript.

Response: Thank you for pointing out. We did not use this consistently as it is also used inconsistently in the literature We checked the text and used exposure-response for all environmental

measurements related to human responses. However, the measure of Mn in body fluids is not equivalent to dose in the toxicological of pharmacological sense. This is defined as the total amount of an agent administered to or taken up or absorbed by an organism, system, or (sub)population [1]. In most cases –dose- is used as a proxy for the inhaled exposure and in fact a measure of its bioavailability after metabolism. We used dose in the sense of the measures (whether a single average or a composite with intensity and duration) of inhaled Mn. To avoid confusion, we changed dose in the text to exposure dose.

- the results section should present some quantitative statements as to reported associations, preferable stratified by important health outcomes.
- since the conclusion suggests better exposure definition and assessment, I would expect some information in the results section that supports this advice, since otherwise it will remain a too generic statement.

Response: Thank you for your feedback we have tried to be more specific in the said sections now

2. Background:
   - P7: it is reported that cognitive function may improve after exposure cessation, but no information is presented about the time window. This is highly relevant, as this will give some guidance for appropriate biomonitoring strategies.

Response: Thank you. We agree that it would be crucial to know a time window in which a removal or reduction of exposure would result in reversal of symptoms, however no evidence of the nature (duration) of such window has been reported specifically in primary studies (cited) therefore we refrained from extrapolating. This has been added to the text.

   - P7: several guidelines are presented, but I would like to read also the critical health outcomes that underlie these risk assessments.

Response: We did mention that these were neurological function deficits, however, the guidelines themselves were not specific about a disease or outcome.

   - P7: see also above. The last sentence is very confusing, is it about the association between air measurements and blood levels (which is not dose-response), or about differences between exposure-response and dose-response associations?

Response: thanks for pointing it out, corrected

   - P8: it is not very clear why this particular review is needed when there are apparently several systematic reviews and meta-analyses already available.

Response: The first aim is to provide an overview of existing evidence based on systematic reviews. Because we found that searches were limited in the existing reviews and the evidence presented was old as well as selective. Furthermore, the evidence was not appraised for quality so as to inform the conclusions. Last but not least, reviews of evidence need updating at every 2 to 5 year intervals and we approached this review as an update, to begin with, but found that a new review was needed as the previous ones were not comprehensive.

3. Methods:
   - P9: the description of the empirical studies, esp the case-control studies and timing of exposure information remains unclear. Does this imply that only nested case-control studies are eligible for inclusion, or that exposure information must have been ascertained through an independent source available prior to diagnosis ?

[1] WHO. PRINCIPLES FOR MODELLING DOSE–RESPONSE FOR THE RISK ASSESSMENT OF CHEMICALS Environmental Health Criteria 239, WHO, Geneva 2009

Response: We include any studies that provided a change in the outcome from start of exposure over any period of time. Exposure did not have to be an independent source but yes, must be done before start of follow up.

- P10: there is no information on the period covered by this search strategy.

Response: Thank you, corrected now

- P11: the assessment of bias is presented, but for exposure assessment there is some need to link the issues rated to the mechanisms of attenuation of exposure-response associations due to random exposure misclassification. For example, was an individual or group-based strategy used to characterize exposure, how well do the measurements of exposure reflect the true exposure at individual level or group level. Maybe this should be limited to the discussion, but it is important to reflect on measurement strategies in more detail somewhere in this report, since the classical topics of bias do not address very well the profound impact of choices in measurement strategies.

Response: Thank you for your detailed comment. We inferred from the included studies that we need information on at least the following three dimensions: time covered (hours, years), intensity (mm3), and particle form or size. In addition, it would be good to know more about the effects of specific exposure measurement strategies. We have added this to the discussion.

- P11-12: since both reviews and original studies are included, I was expecting some statements about potential overlap between these two sources.

Response: We believe that this is not very informative. Reviews were sought so that we won't need to do the next step, once we say we need to do a review ourselves it is based on the finding that the existing ones are not sufficient, then any overlap is not relevant, but only proof of what we found: these reviews were not comprehensive. Also, an overlap would be important if the materials and methods used in reviews were the same, which was not the case.

- P12: the statement that there is no clear pattern between increasing exposure and development of symptoms is unclear. Which dimension of exposure is at stake (duration, frequency or magnitude?). With respect to my earlier remark, what is the relevant time-window here?

Response: We are not sure if this mean years in duration or shift length duration? In any case, we reiterate that this is an area needing expert consensus urgently to ensure uniformity of the measuring exposure and its outcomes in future studies. In the absence of clear definitions of exposure and what aspects or dimensions it constitutes prevents drawing clear conclusions.

- P12: data synthesis: this requires a more precise description for various reasons:
  * a CC study cannot present a RR, so was this limited to cohort studies? (or OR interpreted as an RR)

Response: Odds ratios were interpreted as risk ratios if the incidence of the condition was less than 10% which was the case for all health outcomes. This is added to the text

  * an RR will be available from every cohort study, and, thus, under certain assumptions, it should be possible to present for each study an RR for the same exposure unit (as average exposure is usually reported). A classic example of this approach is the Hodgson and Darnton review on asbestos from 2000.

Response: We are not sure if we understand. Does this mean that we should have done a meta-regression based on average dose per study? If so, we believe that this is not very reliable in case of very few studies because it would be a regression analysis with for example five cases. The RR per unit

of exposure could not be calculated because we didn't have increasing exposure response evaluated in studies.

* when calculating a pooled RR, independent of underlying differences in for example average exposure in the cohort or contrast in exposure, essential exposure-response information may be masked completely.

Response: hence our avoidance of pooling in all except one outcome.

- P13: the subgroup analysis could not be carried out due to lack of data. Is this a correct statement, I think it is rather lack of enough studies for a meaningful analysis.

Response: changed to studies now.

4. Results:
   - The flow chart shows that 6 out of 8 original studies were not included in the systematic reviews. This is a truly interesting finding. Does this indicate that most systematic reviews were conducted on non-informative studies?

Response: Yes.

   - The row with air exposure information in table 1 is too difficult to interpret, since several definitions are included which hampers interpretation, eg how to interpret Boojar 2002?

Response: Boojar clearly has an exposure that would be unlawful in most western countries (and yet we see no motor deficit reported). We think it is quite clear how different this study is from others and how similar the others are in actual exposure measures while reporting variable outcomes. The row also shows that only two studies reported the year of measurement. So we really know very little even if it seems a lot.

   - P21: risk of bias: I am not sure I agree with the claim that review quality is determined mostly by these 3 aspects. This may be true for the reproducibility of the review, but for overall interpretation I worry much more about bias in exposure and outcome assessment, since that will impact the exposure-response association.

Response: Risk of bias in reviews was not assessed on any three aspects but on a validated scale called AMSTAR. These three are not risk of bias items but the two major types of reviews that we came across discussed separately: those of nature of exposure, and those on adverse health effects. The overall risk of bias in the entire set of reviews is then presented as a whole.

   - P21: Last paragraphs: This seems like cutting corners. A cross-sectional study can deliver very informative information when the time window between exposure and health effect is short. Thus, this remark must be substantiated for health outcome of interest.

Response: These reviews were not about health outcomes but about the range and nature of exposure. This is why we think the cross sectional studies were appropriate in these.

   - P21: the statement of Baker remains unclear. Is there an apparent lack of association between air measurements and biomarkers of exposure? Are biomarkers of exposure considered superior ? (When applicable, I am not sure about such a statement without any information of half-time and variability patterns in human material and in air).

Response: We agree with you completely. However, please note that Baker only identified the limitations of current evidence and proceeded to elaborate methods of an ongoing study to address these very doubts.

   - P24: the statement that "the absence of protocols for risk factor studies further limits…" is rather intriguing since there are several instruments published in recent years and their applications have shown their usefulness. My favourite is certainly:

Vlaanderen J, Vermeulen R, Heederik D, Kromhout H. 2008. Guidelines to evaluate human observational studies for quantitative risk assessment. Environ Health Perspect 116:1700–1705.

Hence, a more modest statement is required here.

Response: It seems we have been misunderstood here. Protocols mean the prior publication of the methods for a study either ongoing or about to be undertaken (as Baker et al do in their review). This is not about risk of bias instruments. Protocols prevent data driven findings and allow clear and accurate assessment of risks of bias in the evidence and are therefore a requirement for clinical comparative studies across the world now. We think a similar initiative for all types of studies is needed. Thank you for referring to the Vlaanderen guideline. We believe that this has been superseded by the guidelines presented by Woodruff and Sutton Environmental Health Perspectives 2014 that are much better underpinned than the Vlaanderen guidelines.l

- P25: Some remarks about alternatives for matching must be presented with less vigour, since matching will be extremely difficult for many reasons in occupational groups. Also, studying the mean can be very informative, see Hodgson et al 2000, when evidence is compiled across studies rather than within studies.

Response: Internal validity of a study is important and if we are claiming that one level or type of exposure is dangerous compared to no exposure we are essentially claiming causation. It is essential that such a claim is substantiated with rigorous methods to prevent bias due to misclassification and confounding. Matching prevents it whereas post hoc analyses only adjust for those that were found significant.

- P28: have these studies on association between air and biomarkers adjusted for potential confounders? It seems that if diet is very important, is should be given consideration.

Response: We based our selection of confounders on team consensus after appraising available primary evidence on each, in the absence of evidence based pre-existing expert consensus on any of these. Indeed, these need to be tested for how much weight each of these may carry. Again, protocols developed a-priori for primary studies would help establish these faster and more accurately.

- P32: I do not understand the definition of the fertility outcome on frequency or number of conceptions, since these measures do not reflect (in)fertility, which would be something like delayed time to pregnancy.

Response: We only reported what we found in the included review. There were no studies that reported time-to-conception or to-pregnancy. There were neither studies that reported semen quality or sperm count/ health. This has now been added to the text.

- P33: Table 5 and the underlying review are difficult to interpret, since what is the exposure measure of the presented RR. What is the actual comparison (welders/cutters against other occupational groups?)

Response: This is exactly what is presented: a comparison of welders/cutters against other groups, which is undefined. This should make clear our argument about how uninformative these reviews are, although often there is a lack of information in primary studies, the reviews do not start by clearly defining what comparisons are to be addressed and what is the exposure definition and control definition.

- P34: see earlier remark, I cannot interpret the air exposure row (holds for several tables). Also, what is the comparison in both studies? Eg. Park has a RR of exposed vs referent, but was referent the very low category?

Response: Please see our previous response about air exposure levels. No the referent category is a no exposure group in Park. Although no measurements for that group are provided.

- P35: are these weights correct, since applying these weights to the RR presented will definitely give you a higher value than 1.72.

Response: Yes, the calculations are correct. Please, note that these are not the weights used for the calculations. Since these is a random effects model the weight per study is 1/ within study variance + 1/ between study variance. This is the factor used to multiply the log RR with. Finally, in this model, the first study gets 62% of the weight and the second study 38%. Please, also note that the fixed effect model would have very different results

- P40: after reading all the details for each neurological outcome, I lost the general picture. It would help to include some summarizing tables of statements made by review authors on associations and confront these with your own evaluation, maybe even across all four groups of health outcomes as part of the discussion.

Response: We believe we did that in the conclusions. A table of the same is presented in GRADE format for easier and quick interpretation.

5. Discussion:

- P45, summary statement: see summary remarks. I would like to see a more quantitative description here of the main results, also reflecting the four disease categories studied.

Response: Unfortunately, we do not have enough quantitative data to report. The findings were heterogeneous and therefore it is better to stay away from quantities because these may be misinterpreted.

- P45, nature of exposure: be careful here, since there are certainly large-scale exposure studies with this information, but these may not have been retrieved in the search strategy. Some occupational hygiene journals have published detailed workplaces studies. It is probably true that existing reviews rely only on information presented in the epidemiological studies, whereas most often the exposure surveys are published somewhere else.

Response: Our search was comprehensive including not only peer reviewed sources but also grey literature from many electronic and non-electronic sources. We appraised all surveys however these did not fit the inclusion criteria for various reasons which are presented in the excluded studies table. We further elaborate on the cross sectional studies that did not get to a full text appraisal in our discussion.

Finally, a survey would tell us what the average exposure is for a certain population at a certain point in time and what the variations of this exposure are within this population. Unless these surveys are combined in a systematic way, considering the time windows, we won't know what the range of exposure in air is for a wide range of work places. This is important work and must be undertaken.

- P46: the biomarker discussion…there is ample evidence that some markers of exposure do not reflect exposure very well and thus may attenuate completely the true risk, whereas the air measurements will show a clear association (I recall an example of Rappaport on I think styrene exposure and DNA adducts warning for the over reliance of biomarkers). Thus, I suggest to phrase this more carefully, since an association among these three measures is *not* required.

Response: We agree. A clear statement from experts to prevent the use of such measures which make poor proxies in future studies is needed. Otherwise, a clear proof of an association or a lack of association between the two measures is needed.

- P46: these two studies on PD do not present opposite results, but rather different results!

Response: changed

- P51: the discussion on better exposure characterization is welcomed, but some statements seem to be personal opinion rather the evidence-based. For example, why should we embark on different guidelines for different particle sizes rather than defining the most critical particle size with respect to health? [see the asbestos history, some attention for particle size distributions in exposure was not very helpful for adequate protection of the workforce, to say it politely)

Response: We agree. There is no evidence so we can only extrapolate from what is available. We have now added this suggestion to the text.

- With respect to PD I was expecting some considerations on the importance of smoking as protective factor and inability of most studies to adequately adjust for smoking status.

Response: We expected it too, but did not find enough information. We think it gets covered in confounding as a whole, which a lot of studies do not consider when starting a follow up of a cohort.

Copenhagen 17, 2014

Independent external evaluation of a review commissioned by the Danish Occupational Health Research Fund:

Review of occupational exposure to manganese and the potential health effects of such exposure

Written by

Sharea Ijaz

Jos Verbeek

Anneli Ojajärvi

Markku Sainio

Grant McMillan

Cochrane Occupational Safety and Health Review Group, The Finnish Institute of Occupational Health

Helsinki, Finland

Summary

The purpose of the review was to evaluate the evidence for adverse health effects of occupational exposure to manganese, to examine dose-response relationships and to define a no observed adverse effect level (if any).

The evaluation is based upon thirteen systematic reviews and eight primary studies - all identified by specified inclusion and exclusion criteria from multiple electronic database searches supplemented with reference screening. The thirteen systematic reviews were addressing occupational manganese exposure and different adverse health outcomes regardless of study design of included studies, while the eight primary studies only included prospective and retrospective follow-up studies comparing newly developed outcomes in workers exposed to manganese with less exposed workers. Outcomes included neurologic disease such as Parkinson's disease and Amyotrophic Lateral Sclerosis (ALS), functional neurological test outcomes, respiratory disease, cancer and fertility.

It is concluded that there is insufficient evidence to establish causal links between occupational manganese exposure and any adverse health effect. The database did not allow for any conclusions with respect to existence of exposure-response relationships. Moreover, a no observed effect level could not be established from the reviewed literature.

Critical comments

The review is comprehensive. The authors have managed to describe and summarize information from a very broad and inhomogeneous research field including many different adverse health effects. The authors adhere rigorously to guidelines for evaluation of systematic reviews and for evaluation of the scientific quality of primary research articles. Study selection, data extraction and assessment of potential bias was performed independently by two from the research group including professor Jos Verbeek, who is internationally acknowledged for his work to promote the quality of scientific reviews. In the following I mention three main concerns. In the attached copy I have provided 22 specific comments/questions to specific parts of the report.

1. Inclusion criteria

The criteria set up to select appropriate reviews and primary papers for the review might be described more clearly in the abstract and text, please see the attached copy. In particular I am in doubt about the criteria with respect to exposure. Was it requested that manganese somehow was measured/quantified which would mean that a large number of studies addressing metal welders would escape evaluation even welders have a well characterised exposure to manganese particulates which even may be estimated if total fume exposures have been measured or approximated. If so several large and informative follow-up studies of adverse health outcomes in relation to manganese exposure have been omitted.

Response:

Thank you for your comments and appreciation of the review. The inclusion was restricted to any studies that provided any quantitative assessment of exposure to manganese in a working population. We do not think that studies ascribing exposure to be a job title are informative for deciding about a dangerous level of exposure for worker safety and its resultant effects. We also disagree that welders' exposure to Manganese can be easily deducted from other sources. This might be possible in some exceptional cases where the welding materials have been extensively described. However, in the majority of welding studies, there is hardly any information on the type of material that the welders use. This means that there is a range of metals that they can be exposed to. This will not help in establishing levels of exposure that lead to adverse health effects due to Manganese exposure.

From the text of the review:

Exclusion of JEM based studies seems an undue conservative approach – JEM studies may provide unbiased information on risk

Response: We do not think that JEM based studies provide unbiased information on exposure because many assumptions are made about job titles reflecting exposure. At a point in history when no better methods of exposure assessment are available the JEM can provide the best information on exposure. However now, when accurate methods are available, we think that these proxy assessments prevent us from finding out the exact concentrations at which the adverse effects start

2. Omission of cross-sectional studies of neuropsychological outcomes

The conclusions reached by the authors, - essentially that the evidence is inconclusive – are based upon a limited number of systematic reviews and primary studies, and therefore one crucial issue is whether all relevant information has been included. I am not aware of any single large piece of scientific work that – if included – would have changed the conclusions substantially. However, while cross sectional studies were included in some of the systematic reviews, studies that were cross-sectional in nature were not included in the selection of primary studies. It is argued that there is general consensus, that it is not possible to draw causal inferences from cross-sectional studies as there is no temporality in such studies. Moreover, that results of causal associations are likely to be biased due to healthy worker effects, participant selection and lack of exposure information in the past. I fully agree on this, when the outcome under study are potentially debilitating diseases such as Parkinson's disease, ALS, respiratory tract disease and cancer, but it does not readily apply to studies of functional neurological tests – and here there is a very large database that has not been evaluated from primary papers. I am not convinced that the decision to exclude primary cross-sectional studies on manganese exposure and functional neurological outcomes is justified. In fact, follow-up studies addressing these outcomes also have inherent problems in terms of timing of the exposure-outcome relation. The ideal study will not only examine change in functional status of the various endpoints, but would also need a shift in exposure status – for instance by examination of newly hired workers – a design which is extremely difficult to implement in practice. The omission of primary controlled cross-sectional studies of neurological functional tests in workers exposed to manganese is in my opinion not justified – in particular not when the included evidence is inconclusive. A key issue in evaluation of cross-sectional findings would be check of consistency across studies with respect to results of the many different tests that have been performed.

Response:

In the case of Manganese, the problem is that there is an abundance of information on Manganese and its possible health effects. The challenge is how to filter out the information that is of sufficient quality to draw valid conclusions. We don't think that we were too strict since we included a substantial number of

systematic reviews even though we could not rate their quality very high. In addition, the quality of the included primary studies was not very high. We do not think that further lowering our quality standards would have led to results that are more conclusive. This is supported by the fact that our results are not really different from the results of the included systematic reviews even though the studies we base our conclusions on are different from most of the reviews.

There are several important problems with neurological tests as health outcomes. Single test outcomes cannot be equated to a diagnosis of a neurological disease and a clinical evaluation is always necessary for evaluating a patient suspected of having neurological disease. We find the naming of the test results such as neurological insult or neurological compromise or neurological insult misleading. Furthermore, for most of the tests, the ranges of normal values are not clear. This makes it difficult to qualify a test result as abnormal. Most tests show differences between genders and ages, as mentioned in the discussion. This means that a slight difference between exposed and control group in age or educational level can already confound the findings. Originally, the tests were meant as a method of health surveillance, diagnosing early effects of exposure before disease has occurred. This has slowly progressed into diagnosing neurological disorder even though there is no evidence that those with low scores on tests will progress to have a neurological disease. These issues make the tests a poor alternative for a real diagnosable outcome, especially when the relied upon measure is the difference between two groups which is not inclusive of change from baseline. Especially in cross-sectional studies, there can be many confounders and selection mechanisms at work that would lead to differences in test results that would be incorrectly ascribed to the exposure.

3. Use of quality scores

The authors discuss comprehensively risk of bias in the various types of studies. With this approach (which in general has been endorsed by many scientific journals) scientific quality according to formal rating of a number of specified study characteristics (for instance exposure and outcome ascertainment) is equivalent to risk of bias. Is this reasonable? Influential epidemiologists as Sander Greenland discourage use of quality scores because 'quality scoring submerges important information by combining disparate study features into a single score' (Rothman et al. Modern Epidemiology, p 681). Moreover, this very systematic approach is to some extent weighing different epidemiological design options equally, and does not provide any indication of the direction of the proposed bias. Some researchers would favour a systematic evaluation of the evidence database by the development of criteria for potential bias and confounding that to fit the particular research field. I do acknowledge, however, that the authors emphasize the poor quality of exposure assessment in many studies in the field as compared to rather valid outcome ascertainment in many studies.

Response:

Thank you for your positive comments on the risk of bias assessment. We fully agree that simply adding up quality scores to one summary score make a poor judgement on risk of bias in primary studies as these miss on what is important for a particular review question. Therefore, we followed the most recommended and current method of assessment as advocated by the Cochrane Collaboration and which is a domain based risk of bias assessment in primary studies. We did not add up the risk of bias score but we defined which domains are most important for this specific study question and based on the risk of bias in these domains defined which studies were judged to be at high risk of bias.

Finally, the quality of all the available evidence was assessed according to the GRADE approach including more items than just quality or risk of bias and it was presented in a table for easy use in decision-making. Therefore, we think that we included a valid assessment of the risk of bias in the included studies and the final judgement of the quality of the evidence. See also Woodruff and Suton 2014 for a more general underpinning of this approach for occupational and environmental science.

In conclusion, the paper provides an excellent systematic and critical review of the limited evidence on adverse health effects of current low-level occupational exposure to manganese, and it is stressed that although there is no consistent evidence of adverse health effect at these exposure levels, this is not equivalent to absence of effects. By and large the evidence is inconclusive. In light of this uninformative result, the main criticism is omission of a large body of cross-sectional studies of neurological functional test in workers exposed to manganese.

Response:

Thank you for your appreciation of our work. It is indeed unfortunate that studies that were performed without measuring a change from baseline were excluded and thus a large body of research was found not informative. The restriction to only high quality evidence is the method of choice for systematic reviews of causal inference so that reliable results can be obtained for decision making in health. Including these one-point poorer value data would have only lowered the validity of our findings, and would not have changed conclusions as you kindly suggested earlier in your comments.

It is important that in future only designs that can answer important questions and studies that test predefined hypotheses are carried out. This will enable scientific growth in the field, prevent duplication of effort as well as waste of resources, and help find answers to many yet unclear areas in worker safety.

Bispebjerg September 26 2014

Jens Peter Bonde

Head of department, professor, Dr Med Sc
Department of Occupational and Environmental Medicine