Long working hours and depressive symptoms: systematic review and meta-analysis of published studies and unpublished individual participant data


This study was the first large-scale meta-analysis of prospective studies on long working hours and the onset of depression, including both published data and unpublished individual participant data (189 729 participants from Europe, Asia, North America and Australia). The findings suggested an increased risk of depression linked to prior overtime work in Europe and, to an even higher extent, in Asia.

Affiliation: Finnish Institute of Occupational Health, PO Box 40, 00032 Työterveyslaitos, Finland. marianna.virtanen@ttl.fi


Key terms: depression; depressive symptom; mental health; meta-analysis; overtime; participant data; psychological distress; systematic review; working hour; working life; working time

This article in PubMed: www.ncbi.nlm.nih.gov/pubmed/29423526

Additional material
Please note that there is additional material available belonging to this article on the Scandinavian Journal of Work, Environment & Health -website.
Long working hours and depressive symptoms: systematic review and meta-analysis of published studies and unpublished individual participant data

by Marianna Virtanen, PhD,1,2 Markus Jokela, PhD,2 Ida EH Madsen, PhD,4 Linda L Magnusson Hanson, PhD,5 Tea Lallukka, PhD,1,6 Solja T Nyberg, PhD,6 Lars Alfredsson, PhD,7,8 G David Batty, DSc,9 Jakob B Bjørner, MD, PhD,4 Marianne Borritz, PhD,10 Herrmann Burr, PhD,11 Nico Dragano, PhD,12 Raimund Erbel, MD, PhD,13 Jane E Ferrie, PhD,8,14 Katriina Heikkinen, PhD,15 Anders Knutsson, PhD,16 Markku Koskenvuo, MD, PhD,6 Eero Lahelma, PhD,6 Martin L Nielsen, PhD,17 Tuula Oksanen, PhD,7 Jan H Pejtersen, PhD,18 Jaana Pentti, PhD,6 Ossi Rahkonen, PhD,6 Reiner Rügulies, PhD,6,19 Paula Salo, PhD,1,20 Jürgen Schupp, PhD,21,22 Martin J Shipley, PhD,6 Johannes Siegrist, PhD,12 Archana Singh-Manoux, PhD,6,23 Sakari B Suominen, PhD,24,25,26 Töres Theorell, MD, PhD,6 Jussi Vahtera, MD, PhD,24,27 Gert G Wagner, PhD,21,28,29 Jian Li Wang, PhD,30 Vasootana Yiengprugsawan, PhD,31 Hugo Westerlund, PhD,7 Mika Kivimäki, PhD1,6,9


Objectives This systematic review and meta-analysis combined published study-level data and unpublished individual-participant data with the aim of quantifying the relation between long working hours and the onset of depressive symptoms.

Methods We searched PubMed and Embase for published prospective cohort studies and included available cohorts with unpublished individual-participant data. We used a random-effects meta-analysis to calculate summary estimates across studies.

Results We identified ten published cohort studies and included unpublished individual-participant data from 18 studies. In the majority of cohorts, long working hours was defined as working ≥55 hours per week. In multivariable-adjusted meta-analyses of 189,729 participants from 35 countries [96,275 men, 93,454 women, follow-up ranging from 1–5 years, 21,747 new-onset cases], there was an overall association of 1.14 (95% confidence interval (CI) 1.03–1.25) between long working hours and the onset of depressive symptoms, with significant evidence of heterogeneity (I²=45.1%, P=0.004). A strong association between working hours and depressive symptoms was found in Asian countries (1.50, 95% CI 1.13–2.01), a weaker association in Europe (1.11, 95% CI 1.00–1.22), and no association in North America (0.97, 95% CI 0.70–1.34) or Australia (0.95, 95% CI 0.70–1.29). Differences by other characteristics were small.

Conclusions This observational evidence suggests a moderate association between long working hours and onset of depressive symptoms in Asia and a small association in Europe.

Key terms depression; mental health; overtime; psychological distress; working life; working time.
Long working hours and depressive symptoms

Depression is a leading cause of years lived with disability, contributing to a significant proportion of disease burden worldwide (1). Given that the burden of mental disorders peaks at working age (1), working-age populations are an important target for prevention. There is growing evidence to suggest a link between work characteristics and the onset of depression, with perceived psychosocial work stress being the most often investigated work exposure (2–7).

Recently, studies have also focused on long working hours as a potential risk factor for mental disorders (8). However, although several reviews of this field exist (5, 9–24), the few systematic quantifications of the evidence have been based on published cross-sectional (18, 19, 25) or published longitudinal studies (26). Given the potential publication bias in studies based on published data (27), an individual-participant meta-analysis of unpublished data would provide important complementary evidence to evaluate the effect of long working hours on mental health. Furthermore, long working hours are more common in Asia and North America than in Europe and the risk of mental health problems is higher among women and individuals with low socioeconomic position than men and those with high socioeconomic status. As determinants and mechanisms that harm and protect mental health may also vary by region and sociodemographic factors, research should examine whether there are systematic differences in the association between long working hours and mental health between countries and employee groups. These issues have not been addressed in previous reviews on long working hours and mental health.

With large prospective data including cohort studies with both published and unpublished data, this review is the first large-scale study of the association between long working hours and onset of depressive symptoms.

Methods

Search strategy and selection criteria

Following the PRISMA guidelines (28), we carried out systematic searches of the PubMed and Embase databases, taking into account publications up to January 2017, using the following search terms without restrictions: (“work hours”, “working hours”, “overtime”) and (“mental health”, “mental disorders”, “psychological distress”, “depressive symptoms”, “depression”). We also examined the reference lists of reviews in the field as well as those of the eligible publications, and performed a cited-reference search of these using Institute of Scientific Information Web of Science (to January 2017). Two authors independently reviewed titles and abstracts of the articles and, for selected publications, carried out fulltext review against inclusion criteria.

The following eligibility criteria were applied to published studies: at least results (eg, abstract, tables) published in English; prospective cohort design with individual level exposure and outcome data with follow-up of 1–5 years; the effect of long working hours examined using shorter hours as a reference group; outcome being onset of depressive disorder, depressive symptoms, depressive disorder, or psychological distress; and results reporting either as relative risk (RR), odds
ratios (OR), or hazard ratios (HR) with 95% confidence intervals (CI), or sufficient information provided to calculate these estimates.

We extracted the following information from each eligible study: name of the first author; year of study entry (publication year if not reported); study site (country); population; number of participants; proportion of women; mean age or age range at baseline; mean follow-up time; measure of depressive symptoms; number and proportion (of study population) of new cases; covariates included in the multivariable models; and the effect estimate for the association between long working hours and depressive symptoms.

Unpublished individual participant data

The second part of the study included unpublished individual-level data from studies included in the Individual-Participant-Data Meta-analysis in Working Populations (IPD-Work) Consortium, which has previously published several meta-analyses on other outcomes (29, 30), or were open access at one of two potential sources: the Inter-University Consortium for Political and Social Research (ICPSR; www.icpsr.umich.edu/icpsrweb/ICPSR) or the UK Data Service (ukdataservice.ac.uk). Each study has existing approval from their relevant local or national ethics committee, and all participants had given informed consent.

Quality assessment

Two authors assessed study quality using Cochrane’s “Tool to Assess Risk of Bias in Cohort Studies” in the following domains: (i) Can we be confident in the assessment of exposure (ie, the predictor variables)? (ii) Were the exposed and non-exposed cohorts selected from the same population? (iii) Can we be confident that the outcome of interest was not present at the start of the study? (iv) Did the statistical analysis adjust for the confounding variables? (v) Can we be confident in the assessment of the presence or absence of confounding factors? (vi) Can we be confident in the assessment of the outcome? (vii) Was the follow-up of the cohorts adequate?

The studies were evaluated in relation to each question using a four-level scale: “definitely yes”, “probably/mostly yes”, “probably/no”, and “definitely no”. The quality of the study was considered high/acceptable if all domains were evaluated favorably, ie, “definitely yes” or “probably/mostly yes” (see details in supplementary table A, www.sjweh.fi/show_abstract.php?abstract_id=3712).

Statistical analysis

We used a two-stage meta-analysis (31) in which study-specific estimates were obtained from published studies and studies with individual participant data at the first stage. In all studies, the baseline cases were excluded in order to allow assessment of onset of depressive symptoms to be examined. With an expectation of heterogeneity between studies, we used random-effects meta-analysis in the second stage to obtain a pooled estimate from the first stage study-specific analyses. We used OR as an estimator of effect and their 95% CI as an indicator of precision, comparing long working hours (most often defined as ≥55 weekly hours) with standard hours (most often 35–40 hours). In cohort studies with individual participant data, the OR were adjusted for age, sex, socioeconomic status, and marital status. For published studies, we used models that were closest to that set of adjustments. Heterogeneity of the estimates was first assessed using the I² statistic. Heterogeneity test and meta-regression were used to assess subgroup differences. Subgroup analyses included sex (men versus women), age group (<50 years versus ≥50 years), socioeconomic status (high, intermediate, low), geographic region (Europe, Asia, North America, Australia), publication status (published versus unpublished), study baseline year (1991–99 versus 2000–10), study population (population based versus occupational cohort), study quality (low versus high/acceptable), length of follow-up (1–2 versus 3–5 years), response rate at baseline (≤65 versus >65%), loss to follow-up (≤25 versus >25%), outcome type (psychological distress versus depressive symptoms), population prevalence of symptoms at baseline (<15 versus ≥15%), and population onset of symptoms at follow-up (<15 versus ≥15%). We tested for publication bias in published studies by using the funnel plot of the estimates against their standard errors and Egger’s test for small-study effects (32). We used SAS 9.4 (SAS Institute, Cary, NC, USA) and Stata 13.1 (StataCorp, College Station, TX, USA) to analyze study-specific data and Stata 13.1 for meta-analyses.

Results

Literature search results

By January 2017, we identified 3295 studies from Embase and 1740 studies from PubMed (supplementary figure A, www.sjweh.fi/show_abstract.php?abstract_id=3712). After screening the abstracts, full text review was performed for 170 articles. Of these, six studies met the inclusion criteria and a further four were identified from referred publication searches, hand search, and from the reference lists of published reviews. We excluded meta-analyses (but extracted relevant studies from those) (18, 19, 25, 26), reviews (5, 8–17, 21, 23), and books (20,
22, 24). We also excluded studies with overlapping data with selective studies (33–40), those without adjustment for baseline mental health (41–43), and those prospective studies in which exposure and outcome were temporally overlapping (44, 45). However, we performed a descriptive, qualitative analysis of studies which included prospective analysis with continuous outcomes rather than outcomes indicating onset of illness (46, 47), a study with exposure treated as a continuous variable (48), longitudinal within-subject design with continuous outcomes (49), and broader mental health outcomes, such as antidepressant use (50, 51), all treated mental disorders (52), or disability claims due to all mental disorders (53). Results of these studies are presented in supplementary table B, www.sjweh.fi/show_abstract.php?abstract_id=3712. None of these studies suggested an association between working hours and mental health.

Descriptive statistics for the identified eligible ten published studies (54–63) are presented in supplementary table C, www.sjweh.fi/show_abstract.php?abstract_id=3712. The range of study entry year varied between 1992–2010. Four studies were from Japan (54, 58, 59, 62), two from Canada (56, 61), and one each from the UK (57), Sweden (55), France (60), and South Korea (63). In three studies, there were separate estimates for men and women (54–56). Adjustment for covariates varied between none (58), stratification by sex (54, 55), and multivariable (56, 57, 59–63).

Unpublished individual participant data
A total of 18 studies (64–81) from the IPD-Work Consortium, open access databases and the Thai Cohort Study (TCS) (80) via personal communication provided unpublished individual-level data for working hours and onset of depressive symptoms (supplementary table C). The study baseline year ranged from 1986–2009. Four studies were from the UK (65, 69, 73, 81), three from the US (64, 66, 67), three from Finland (68, 71, 74), two from Denmark (70, 72), two from Germany (77, 79), and one each from Australia (76), Thailand (80), and Sweden (78). A total of 25 European countries and Israel participated in the Survey of Health, Ageing and Retirement in Europe (SHARE) (75), including France, Germany, Finland, Sweden, Denmark, Slovenia, Cyprus, Romania, Bulgaria, Greece, Czech Republic, Slovakia, Hungary, Estonia, Latvia, Lithuania, Poland, The Netherlands, Belgium, Luxembourg, Spain, Portugal, Italy, Switzerland, and Austria.

In total, the number of participants in published studies and individual participant data was 189 729 (96 275 men, 93 454 women) from 35 countries. The number of cases with new-onset depressive symptoms was 21 747 (the average onset 11.5%).

Assessment of long working hours
With previous evidence suggesting that working ≥55 hours per week may be harmful for health (29, 30) and with the European Union Directive recommending a limit of 48 hours per week (82) in the present analyses of studies with individual-level data, we used the following categories of hours worked per week: <35, 35–40 (reference group), 41–48, 49–54 and ≥55 hours. This categorization was used in all unpublished datasets and via personal communication for a published study by Wang et al (61). In the Helsinki Health Study (HHS) (74), study members had responded to an enquiry about time worked using a pre-defined categorization (<30, 30–40, 41–50 and >50 hours per week) which, for the purposes of the present analyses, we used 30–40 hours per week (reference group) with >50 working hours as the highest group. In the published studies, there was variation in how long working hours were defined in the analyses: >12 versus ≤8 hours /day (54); ≥41 hours /week versus less (56); >48 hours a week versus less (60); >60 versus ≤50 hours /week (62); 60 or more hours versus less (59); >55/week versus 35–40 (57); >68 hours versus 35–40 (63); ≥40 hours overtime /month versus less (58); and "often working overtime" versus not (55). To evaluate whether this heterogeneity in exposure assessment had an effect on results, we conducted separate analyses for published studies and unpublished individual participant data, in addition to overall pooled estimates.

Ascertainment of depressive symptoms
Most of the studies assessed self-reported depressive symptoms or psychological distress using survey questionnaires. Structured interview-based methods based on Diagnostic and Statistical Manual of Mental Disorders (DSM-III) or DSM-IV diagnostic criteria were used in three published studies (56, 60, 61) to assess major depressive disorder or major depressive episode (see supplementary table C).

Assessment of individual- and study-level characteristics
The published studies had used different covariate adjustments (supplementary table C). In cohort studies with the individual participant data, we extracted the following: age, sex, marital status (married/cohabited versus not) and socioeconomic status (SES) which, based on register or survey-based occupational position, or educational qualification, was categorized as high, intermediate, or low. Other factors included geographic region (Europe, Asia, North America, Australia), publication status (published versus unpublished), study baseline year (1991–99 versus 2000–10), study population (population based versus occupational cohort), study
quality (low versus high/acceptable), length of follow-up (1–2 versus 3–5 years), response rate at baseline (≤65 versus >65%), loss to follow-up (≤25 versus >25%), outcome type (psychological distress versus depressive symptoms), population prevalence of symptoms at baseline (<15 versus ≥15%), and population onset of symptoms at follow-up (<15 versus ≥15%).

Quality assessment
Results from the study quality assessment are presented in supplementary table A. The scores ranged between 7–11 out of the best possible 14 points; the best (11 points) in the present analysis being assigned to the studies by Shields (56) and Niedhammer et al (60). Of the 28 studies, all but three (54, 55, 58) were considered being of high/acceptable quality, that is, having no components with serious threat of bias. The three studies lost score points due to lack of confounder adjustments and inadequate exposure. To examine potential other sources of bias not assessed in the tool, we included in the subgroup analyses publication status, study baseline year, type of population, response rate at baseline, loss to follow-up, outcome type, prevalence of depressive symptoms in the study population at baseline and proportion of those with new-onset of symptoms at follow-up (see also supplementary table D, www.sjweh.fi/show_abstract.php?abstract_id=3712).

Long working hours and onset of depressive symptoms
Study-specific association between long working hours and onset of depressive symptoms is presented in figure 1. The overall OR was 1.14 (95% CI 1.03–1.25) after multivariable adjustments. The F statistics (F=45.1%, P=0.004) suggested significant heterogeneity between the studies.

A significant difference was found for Asia versus other geographic regions (P=0.034 using heterogeneity test; P=0.041 using meta-regression). Also in figure 1, study-specific forest plots by geographic region are shown for the association between long working hours and depressive symptoms. The OR was 1.50 (95% CI 1.13–2.01) for Asia, 1.11 (95% CI 1.004–1.22) for Europe, 0.97 (95% CI 0.70–1.34) for North America, and 0.95 (95% CI 0.70–1.29) for the one study representing Australia. However, significant heterogeneity was still present in subgroups of studies from Asia and North America although all Asian studies suggested a positive association between long working hours and depressive symptoms.

Results from other subgroup analyses are displayed in figure 2. There were two characteristics with evidence of subgroup differences; published versus unpublished data (P-values 0.075 and 0.018 using heterogeneity test and meta-regression, respectively), and population-based versus occupational cohort (P-values 0.048 and 0.092). Published studies indicated a 1.35-fold (95% CI 1.07–1.71) risk of depressive symptoms whereas unpublished data suggested a 1.08-fold (95% CI 1.00–1.16) increased risk. The estimate for occupational cohorts was 1.34 (95% CI 1.10–1.62) and for population-based studies 1.07 (95% CI 0.96–1.20). Meta-regression P-values for other estimates (not shown in the figure) were as follows: study baseline year P=0.458; study quality P=0.440; length of follow-up P=0.958; response rate P=0.805; loss to follow-up P=0.720; outcome type P=0.415; population prevalence of symptoms at baseline P=0.880; population onset of symptoms P=0.308. Furthermore, no significant subgroup differences were found when the outcome type was categorized into three groups: "psychological distress", "depressive symptoms", and "clinical depression" (P=0.619 and 0.588 in heterogeneity test and meta-regression, respectively). The Egger test of publication bias did not show evidence for publication bias neither among the ten published estimates nor among the subgroup of Asian-origin studies. The OR for the association between long working hours and depressive symptoms was 1.65 (95% CI 1.30–2.09) in published studies from Asia whereas it was 1.07 (95% CI 0.74–1.55) in other published studies (P for difference=0.053). Thus, a stronger association in Asia than in other countries was found both in all studies and within published studies.

Discussion
In this systematic review and meta-analysis of nearly 190 000 participants from 28 prospective cohort studies in 35 countries, the overall association between long working hours and depressive symptoms was 1.14-fold (95% CI 1.03–1.25), with significant heterogeneity between studies. The subgroup analyses revealed that the summary estimate appeared to vary by geographic region; whereby the pooled estimates suggested no association between long working hours and depressive symptoms in the studies from North America and Australia and a small association in the studies from Europe, it was larger (OR 1.50, 95% CI 1.30–2.09) in the studies conducted in Asian countries.

Of the three previous meta-analyses of the association between long working hours and mental health, two were published in 1997 (18) and 2008 (19), and were mainly based on cross-sectional data. They reported a weak linear correlation between hours worked and "mental strain", which is consistent with our findings. A more recent meta-analysis identified seven published prospective studies and a summary estimate which
Long working hours and depressive symptoms

indicated an increased (RR=1.08) but statistically non-significant risk of clinical depressive disorder associated with long working hours (26). However, all previous meta-analyses were exclusively based on published data and their ability to examine subgroup differences was limited.

In our meta-analysis, the association between long working hours and depressive symptoms was stronger in Asian countries, including studies from Japan, South Korea and Thailand, than the rest of the countries from Europe, North America, and Australia. The reasons for this regional difference are unclear, but they might involve cultural and occupational health policy differences between Asian and Western societies. According to the effort–reward imbalance model (83), health risks of long working hours are smaller if efforts at work (long hours) are rewarded by economic rewards, promotion, esteem, and high job security. In the ideal scenario, working long hours is voluntary and leads to increased rewards; the unfavorable scenario involves involuntary long working hours combined with low rewards (84). Another theoretical framework, the effort–recovery model (85) posits that long working hours may reduce the time for recovery to an extent that increases the risk of fatigue and associated mental health symptoms. In the present meta-analysis, however, we were not able to determine whether effort–reward or effort–recovery imbalance mediate the association between long working hours and depressive symptoms. However, a study from South-Korea showed that low job security, a component of low rewards, predicted depressive symptoms, suicidal ideation and impaired self-rated health (86). In addition, in Japan, sense of community is a fundamental feature of the society, and an important endeavor in individuals is to be in harmony with the community – including working community – rather than achieving one’s individualistic goals (87). The Asian work culture, described as "worker bees", may lead to a prolonged

Figure 1. Random-effects meta-analysis of the association between long working hours and the onset of depressive symptoms by geographic region. OR, odds ratio.
struggling with overwhelming workload without seeing any other options (88). Suicide (Karojisatsu) caused by overwork has been suggested to be common because showing weakness, such as depression, in work spheres means betraying one’s co-workers (87). Other issues in Asian countries might relate to limited prevention and treatment of depression and social stigma related to mental disorders, which might restrain help-seeking at an early phase of the disease (88).

In addition to geographic region, an effect modifier was found for publication status and cohort type. The estimate suggested a stronger association in published studies than unpublished studies although the funnel plot test suggested sources other than small study publication bias underlay this difference. The association was also stronger in studies that were based on occupational cohorts than in those derived from the general population. However, as there was only one study from Asia in the unpublished group and one in the population-based group, further meta-analyses with a greater number of Asian studies are needed to determine whether these other subgroup differences underlay the geographic difference in the association between long working hours and depressive symptoms.

The studies included in this meta-analysis have some important limitations. Exposure to working hours was based on self-reports in all studies and measured only once. Self-report may involve recall bias due to the participants’ inability to recall hours worked accurately. However, the validity of self-reported working hours has been found to be moderate or good (89, 90). A non-systematic error in the assessment of working hours is likely to attenuate the association between long working hours and depressive symptoms and recall bias may artificially inflate the observed associations only if the error in the number of hours worked is systematically correlating with mental health status. Furthermore, with exposure measured at one time point only, we were not

---

<table>
<thead>
<tr>
<th>Subgroups</th>
<th>OR (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>1.08 (0.97, 1.20)</td>
<td>0.523</td>
</tr>
<tr>
<td>Women</td>
<td>1.14 (1.00, 1.29)</td>
<td></td>
</tr>
<tr>
<td><strong>Age group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50 years</td>
<td>1.19 (1.03, 1.39)</td>
<td>0.223</td>
</tr>
<tr>
<td>≥50 years</td>
<td>1.01 (0.81, 1.25)</td>
<td></td>
</tr>
<tr>
<td><strong>Socioeconomic status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>1.07 (0.84, 1.37)</td>
<td>0.837</td>
</tr>
<tr>
<td>Intermediate</td>
<td>1.13 (0.98, 1.31)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>1.07 (0.85, 1.32)</td>
<td></td>
</tr>
<tr>
<td><strong>Publication status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Published</td>
<td>1.35 (1.07, 1.71)</td>
<td>0.075</td>
</tr>
<tr>
<td>Unpublished</td>
<td>1.08 (1.00, 1.16)</td>
<td></td>
</tr>
<tr>
<td><strong>Study baseline year</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991-99</td>
<td>1.19 (1.04, 1.36)</td>
<td>0.361</td>
</tr>
<tr>
<td>2000-10</td>
<td>1.09 (0.96, 1.25)</td>
<td></td>
</tr>
<tr>
<td><strong>Population-based study</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1.07 (0.96, 1.20)</td>
<td>0.048</td>
</tr>
<tr>
<td>No</td>
<td>1.34 (1.10, 1.62)</td>
<td></td>
</tr>
<tr>
<td><strong>Study quality</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>1.29 (1.00, 1.65)</td>
<td>0.306</td>
</tr>
<tr>
<td>High/acceptable</td>
<td>1.12 (1.01, 1.24)</td>
<td></td>
</tr>
<tr>
<td><strong>Follow-up (years)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>1.14 (0.99, 1.31)</td>
<td>1.000</td>
</tr>
<tr>
<td>3-5</td>
<td>1.14 (1.00, 1.29)</td>
<td></td>
</tr>
<tr>
<td><strong>Response rate at baseline (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤65</td>
<td>1.09 (0.99, 1.21)</td>
<td>0.663</td>
</tr>
<tr>
<td>&gt;65</td>
<td>1.13 (1.00, 1.29)</td>
<td></td>
</tr>
<tr>
<td><strong>Loss to follow-up (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤25</td>
<td>1.12 (1.01, 1.25)</td>
<td>0.573</td>
</tr>
<tr>
<td>&gt;25</td>
<td>1.20 (0.97, 1.48)</td>
<td></td>
</tr>
<tr>
<td><strong>Outcome type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological distress</td>
<td>1.18 (1.06, 1.32)</td>
<td>0.395</td>
</tr>
<tr>
<td>Depression</td>
<td>1.09 (0.94, 1.26)</td>
<td></td>
</tr>
<tr>
<td><strong>Population prevalence of symptoms at baseline (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;15</td>
<td>1.13 (0.92, 1.39)</td>
<td>0.883</td>
</tr>
<tr>
<td>≥15</td>
<td>1.11 (0.98, 1.24)</td>
<td></td>
</tr>
<tr>
<td><strong>Population onset of symptoms at follow-up (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;15</td>
<td>1.08 (0.94, 1.24)</td>
<td>0.253</td>
</tr>
<tr>
<td>≥15</td>
<td>1.20 (1.07, 1.35)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Subgroup analyses for the association between long working hours and the onset of depressive symptoms. [OR=odds ratio.] * P-value for difference between groups.
able to assess whether the relatively small overall risk found in this meta-analysis was an underestimate due to misclassification of the exposure, ie, that for some employees, excessive working hours was a temporary work situation as their working hours declined during the follow-up whereas for some others working hours could have increased. Unlike unpublished individual-level data, the definition of exposure to working hours also varied among published studies.

There was also variation in how depressive symptoms were assessed. Some of the studies used more general outcomes of psychological distress while other studies used more specific measures of depressive symptoms or clinical depressive disorder. However, the association was similar across these three outcome measures. In only three studies (56, 60, 61), structured interview-based methods for DSM-III or DSM-IV diagnostic criteria for depressive disorder was used, thus the majority of studies in this meta-analysis concerned depressive symptoms or psychological distress. However, sub-clinical depression has been shown to be a more important indicator of health than previously thought (91) although the association found in our study was not dependent on the outcome type.

Furthermore, bias due to unmeasured confounding cannot be ruled out in observational studies, which can of course only provide insights into association, not causality. For example, unmeasured work-related (such as low job control, job strain and shift work) and behavioral factors (such as alcohol consumption) might have contributed to the association between long working hours and depressive symptoms. However, previous studies show that long working hours are associated with high job demands and high control ("active job") rather than high strain jobs (high demands with low control) (92, 93), making low job control and job strain unlikely confounders. Similarly, there is no strong longitudinal evidence linking shift work to depression (94), and long working hours has been found to be only modestly associated with risky alcohol use (30). In combination, unmeasured confounders should be associated both with long working hours and depressive symptoms with an odds ratio of 1.54 (95% CI 1.21–1.81) to entirely explain the present findings (95). However, one more limitation of the present study is that the vast majority of studies were based on populations from high-income countries; therefore the findings are not generalizable to low- or middle-income countries.

Concluding remarks

In this meta-analysis of prospective studies with up to 190 000 employees from 35 countries, long working hours, typically defined as ≥55 weekly working hours, were associated with a modest 1.14-fold increased risk of new-onset depressive symptoms, with high level of heterogeneity between studies, but also evidence that this association may be context-specific. Accordingly, the odds of having depressive symptoms at follow-up was 1.5-fold in individuals working long hours in Asian countries while the corresponding OR was 1.1 in Europe and less than that in North America and Australia. Further research is needed to determine the specific reasons behind these differences and to develop policies to prevent depression due to excessive working hours.

Acknowledgements

NordForsk, Academy of Finland (311492, 258598, 292824), Medical Research Council UK (K013351), Finnish Work Environment Fund, and a Helsinki Institute of Life Science Fellowship supported this review.

Conflict of interest

None declared.

References

7. Madsen IE, Nyberg ST, Magnusson Hanson LL, Ferrie JE,


37. Michäelsen H, Bildt C. Psychosocial conditions on and off the job and psychological ill health: depressive symptoms, impaired psychological wellbeing, heavy consumption of alcohol. Occup Environ Med 2003 Jul;60(7):489–96. http://dx.doi.org/10.1136/oem.60.7.489


55. Bildt C, Michäelsen H. Gender differences in the effects from working conditions on mental health: a 4-year follow-up. Int


Received for publication: 20 October 2017