The WHO/ILO report on long working hours and ischaemic heart disease – Conclusions are not supported by the evidence

Mika Kivimäki, Marianna Virtanen, Solja T. Nyberg, G. David Batty

A R T I C L E   I N F O
Handling Editor: Paul Whaley
Keywords: World Health Organization
International Labor Organization
Long working hours
Ischaemic heart disease
Systematic review
Meta-analysis

A B S T R A C T

Working hours is a ubiquitous exposure given that most adults are employed, and one that is modifiable via legislative change if not always through individual-level choice. According to a recent report from the World Health Organization (WHO) and International Labour Organization (ILO), there is currently sufficient evidence to conclude that long working hours (i.e., ≥ 55 h per week) elevate the risk of fatal and non-fatal ischaemic heart disease to a clinically meaningful extent. After assessing the data used by the ILO/WHO, we feel that the expert group has not correctly applied their own framework for assessing the strength of the evidence. In the meta-analysis of observational studies in the report, the association between long working hours and incident heart disease appeared stronger in lower quality cohort studies with a high risk of bias (minimally-adjusted hazard ratio 1.20, 95% CI 1.01–1.41, compared to standard 35–40 weekly hours) than in the superior-quality studies with a lower risk of bias for which the estimate was not significantly different from the null (1.08, 95% CI 0.93–1.25). There was also marked effect modification, such that there was no increase in ischaemic heart disease for those working long hours in high socioeconomic status occupations, a finding also reported in analyses of a recent census-based cohort study which was not included in the report. Our meta-analysis of all these studies confirms that the findings are not consistent but differ between subgroups and that the summary age- and sex-adjusted hazard ratio for long working hours in high socioeconomic status occupations does not support excess risk: 0.85, 95% CI 0.63–1.13 (Pinteraction = 0.005, total N = 451,982). For these and other reasons detailed in this commentary, we advance a more cautious interpretation of the existing evidence. The conclusions should be restricted to low socioeconomic status occupations only and more research is still needed to confirm or refute harmfulness and determine clinical relevance.

The United Nation’s World Health Organization (WHO) provides syntheses on risk factors for disease and injury for its 194 member states. A recent report by WHO and International Labour Organization (ILO) (Li et al., 2020) focused on working hours, a ubiquitous exposure given that most adults are employed, and one that is modifiable via legislative change, if not always through individual-level choice. According to the report there is currently “sufficient evidence of harmfulness” to conclude that long working hours (i.e., ≥ 55 h per week) elevate the risk of fatal and non-fatal ischaemic heart disease and that this excess risk is “clinically meaningful” (p. 1) (Li et al., 2020).

To encourage a wider discussion on the validity of the WHO/ILO’s conclusions, we provide five specific points in support of our view that the report has overstated the strength and generalisability of the available evidence. To support our argument that the association between long working hours and ischaemic heart disease is likely to be dependent on the type of work rather than being universal observation, we present results from a new, updated meta-analysis.

1. WHO/ILO conclusions are not supported by other published data syntheses

Previous meta-analyses suggested the association between long working hours and ischaemic heart disease is not robust. In a 2015 meta-analysis of 22 prospective cohort studies, 598,470 participants and 4652 incident cases, the age-, sex- and socioeconomic status-adjusted summary hazard ratio for individuals working ≥ 55 weekly hours compared to those working standard 35–40 h was a modest, 1.13; 95% CI 1.02–1.26 (Kivimäki et al., 2015). In the multivariable-adjusted model, the summary hazard ratio was attenuated and statistically non-significant at conventional levels: 1.08; 95% CI 0.94–1.24. With the
addition of results from the Danish Labour Force Survey, (Hannerz et al., 2018) an update of this meta-analysis was published in 2018. Comprising 23 cohort studies and a total of 744,331 participants in whom there were 8287 incident heart disease cases, (Virtanen and Kivimäki, 2018) the minimally-adj usted summary hazard ratio was little changed (1.12; 95% CI 1.03–1.21) and the authors concluded that “the possibility of residual confounding and bias cannot be ruled out.” In 2020, Rivera and colleagues produced an umbrella piece – another data synthesis but based on a systematic review of systematic reviews – for long working hours and chronic conditions (Rivera et al., 2020). They concluded that the evidence for ischaemic heart disease was “unclear” and that there is a need for higher-quality studies.

The discordance between the conclusions of these existing reviews (Kivimäki et al., 2015; Virtanen and Kivimäki, 2018; Rivera et al., 2020) and the stronger and more definitive inferences found within the WHO/ILO report (Li et al., 2020) raises the question of what new high-quality data the authors were able to identify. The WHO/ILO main meta-analysis of evidence with acquired ischaemic heart disease as the outcome included 22 cohort studies of which only one (Hayashi et al., 2019) was not included in previous meta-analyses and this study had a low WHO/ILO quality rating whereby the risk of bias was “probably high”. The new study was based on relatively few ischaemic heart disease cases (N = 212) (Hayashi et al., 2019) and therefore the summary hazard ratio in the WHO/ILO meta-analysis was not improved: 1.13; 95% CI 1.02–1.26 in a total of 339,680 participants - the same summary estimate and confidence interval in fact as in the 2015 meta-analysis (Kivimäki et al., 2015).

Thus, the findings of the WHO/ILO review (Li et al., 2020) and other reviews on this topic (Kivimäki et al., 2015; Rivera et al., 2020) are almost identical based on the same pool of studies. This means that the discussion is about the interpretation of the findings. The authors of the WHO/ILO review applied the Navigation Guide as an organising framework (Woodruff and Sutton, 2014) whereas the 2015 meta-analysis (Kivimäki et al., 2015) used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009) and the umbrella review (Rivera et al., 2020) followed the AMSTAR2 (Shea et al., 2017). This raises the question whether different review frameworks were the origin of the discordant conclusions. We believe this was not the case. Instead, our view is that the authors of the WHO/ILO review did not correctly apply their own framework for assessing the strength of the evidence.

2. Subgroup analyses do not support the general conclusion by WHO/ILO

The purpose of subgroup analyses is to investigate whether the effect is consistently observed when the context changes. This constitutes a fundamental step in the assessment of evidence to confirm that the conclusions for the overall study population hold. This was not the case: the relationship between long working hours and ischaemic heart disease was seen in some occupational groups, but not in others. The different associations across the socioeconomic subgroups do not support a general conclusion of harmfulness.

The authors of the WHO/ILO report show large differences in the long working hours- ischaemic heart disease association between socioeconomic groups, the test for heterogeneity for both acquired ischaemic heart disease and deaths from ischaemic heart disease being statistically significant (P-value = 0.04) when a 3-level socioeconomic status variable was used based on occupational group or, when unavailable, educational attainment. Thus, while the association was evident in low socioeconomic status occupations, there was no increase in ischaemic heart disease for those working long hours in high socioeconomic status occupations (hazard ratio 0.87, 95% CI 0.56–1.38).

The WHO/ILO report did not include the findings of a census-based mortality study of more than 400,000 (O’Reilly and Rosato, 2013) because occupational groups were categorised into 4 rather than 3 groups (personal communication with the WHO/ILO group, email 10/06/2020). However, that study replicates the null finding for high socioeconomic status occupations (0.82; 95% CI 0.55–1.23) representing around 40% of the total study population (N = 163,139/414,949).

To provide an overall estimate based on maximum data available, we report in Table 1 an updated socioeconomic status stratified meta-analysis including all cohorts from both the WHO/ILO review and census-based study. The summary age- and sex-adjusted hazard ratio for the association between long versus standard working hours and ischaemic heart disease is 0.85 (95% CI 0.63–1.15, N = 186,079) in the group of employees with high socioeconomic status, 1.23 (95% CI 0.96–1.57, N = 111,672) among those with intermediate socioeconomic status and 1.70 (95% CI 1.27–2.27, N = 154,231) for the low socioeconomic status group (p for socioeconomic status interaction = 0.005, total N = 451,982). In 7 of 9 cohorts, the hazard ratio was below 1 in the groups of people with high socioeconomic status.

Rather than acknowledging the lack of association in the high-socioeconomic status group, the WHO/ILO report notes that relative risk estimates were ‘higher’ among persons with low socioeconomic status. With this evidence, a conclusion which clearly states the significant inconsistency in the association and that the findings do not support long working hours as a risk factor for ischaemic heart disease in individuals with high socioeconomic status would seem more scientifically appropriate.

3. There is not “sufficient evidence of harmfulness”

The Navigation Guide framework defines the criteria for sufficient evidence as: “The available evidence usually includes consistent results from well-designed, well-conducted studies, and the conclusion is unlikely to be strongly affected by the results of future studies. For human evidence a positive relationship is observed between exposure and outcome where chance, bias, and confounding, can be ruled out with reasonable confidence” (p. 16) (NavigationGuide, 2012). If we judge the evidence based on these criteria, the conclusion that there is already “sufficient evidence of harmfulness” cannot be justified.

In addition, a statistically significant but modest summary hazard ratio from non-randomised studies, such as 1.13, raises concerns about residual confounding and bias (Bradford, 1965). Expressed simply, a hazard ratio of such modest magnitude is likely to be pushed towards no effect with the addition of further covariates. Indeed, in the 2015 meta-analysis, additional adjustments reduced effect estimate for long working hours by over 35% and the adjusted association with ischaemic heart disease was non-significant (Kivimäki et al., 2015).

Why then does the WHO/ILO report choose to base their conclusion on minimally-adjusted hazard ratios and ignore those adjusted for a wider range of potential confounders which, in a conventional observational study, are considered the best approximation of the relative risk from a randomised controlled trial? (Lawlor et al., 2004) The justifications for doing so are found in a conceptual model suggesting that other lifestyle factors were consequences of long working hours rather than confounders (Li et al., 2020). Many experts might disagree with this assumption as being overly simplistic. For example, smoking habit which the WHO/ILO report also considered as a mediator is usually adopted before labour market entry and therefore unlikely to be a consequence of long working hours. Indeed, associations robust to multivariable adjustments, such as that between long working hours and stroke [hazard ratio 1.33 (95% CI 1.11–1.61) before, and 1.30 (95% CI 1.05–1.60) after, adjustment for lifestyle factors] (Kivimäki et al., 2015; Descatha et al., 2020) would be more convincing in excluding confounding as an alternative explanation for the findings.

The WHO/ILO report assessed included cohorts for risk of bias, but this did not feature in the interpretations by the authors. In stratified analysis by study quality, the association between long working hours and incident heart disease appeared stronger in cohorts with high risk of bias [hazard ratio 1.20, 95% confidence interval 1.01–1.41,
multivariable adjustments or exclusion of lower-quality evidence, the association between long working hours and ischaemic heart disease was weak and not robust to stratification, differing from the null (1.08, 95% confidence interval 0.93–1.25).

Table 1
Random-effect meta-analysis of long working hours (≥55 h/week) compared to 35–40 weekly hours and ischaemic heart disease by socioeconomic status (SES) in cohorts included in the WHO/ILO report plus a large census-based mortality study.

| Socioeconomic status (SES) group | N (total) | Weight | % | Hazard ratio (95% CD)* |
|----------------------------------|----------|--------|---|-----------------------|
| **Cohort study**                 |          |        |   |                       |
| High SES                         |          |        |   |                       |
| Blue-collar (De Bacquer et al., 2005) | 2303     | 1.9    | 0.77 (0.96–6.89) |
| FPS (Kivimäki et al., 2007)      | 13,241   | 6.1    | 0.97 (0.92–3.26) |
| HeSStu (Korkella et al., 2001)   | 2396     | 2.3    | 2.88 (0.40–20.52) |
| HNR (Stang et al., 2005)         | 830      | 3.5    | 0.32 (0.07–1.57) |
| WOLF-N (Alfredsson et al., 2002) | 363      | 1.2    | 0.95 (0.06–15.28) |
| WOLF-S (Peter et al., 1998)      | 884      | 4.3    | 0.71 (0.17–2.96) |
| Whitehall II (Marmot et al., 1991) | 2923     | 23.7   | 0.92 (0.50–1.70) |
| Census-based, men (O’Reilly and Rosato, 2013) | 94,848 | 54.9 | 0.82 (0.55–1.23) |
| Census-based, women (O’Reilly and Rosato, 2013) | 68,291 | 2.2 | 1.25 (0.16–9.47) |
| **Summary estimate**             | 186,079  | 100    | 0.85 (0.63–1.15) |
| (P = 0%, P = 0.92)               |          |        |   |                       |
| **Intermediate SES**             |          |        |   |                       |
| COPSOQ-I (Kristensen et al., 2005) | 511    | 1.3    | 0.56 (0.07–4.84) |
| DWECS (Feveile et al., 2007)     | 1656     | 4.7    | 0.75 (0.24–2.35) |
| HeSStu (Korkella et al., 2001)   | 23,303   | 4.5    | 1.27 (0.40–4.05) |
| Whitehall II (Marmot et al., 1991) | 3435    | 13.4   | 1.51 (0.77–2.96) |
| Census-based (intermediate), men (O’Reilly and Rosato, 2013) | 45,993 | 52.0 | 1.24 (0.88–1.74) |
| Census-based (own account), men (O’Reilly and Rosato, 2013) | 21,171 | 1.5 | 0.36 (0.05–2.61) |
| Census-based, women (O’Reilly and Rosato, 2013) | 6919  | 18.8   | 1.35 (0.77–2.38) |
| **Summary estimate**             | 111,672  | 100    | 1.23 (0.96–1.57) |
| (P = 0%, P = 0.87)               |          |        |   |                       |
| **Low SES**                      |          |        |   |                       |
| COPSOQ-I (Kristensen et al., 2005) | 776    | 2.0    | 0.85 (0.11–4.68) |
| DWECS (Feveile et al., 2007)     | 1428     | 1.4    | 4.97 (0.44–56.13) |
| HeSStu (Korkella et al., 2001)   | 2356     | 5.6    | 1.41 (0.41–4.79) |
| HNR (Stang et al., 2005)         | 3725     | 6.3    | 2.90 (0.73–7.26) |
| Whitehall II (Marmot et al., 1991) | 330     | 4.2    | 2.80 (0.68–11.54) |
| Census-based, men (O’Reilly and Rosato, 2013) | 1256    | 7.5    | 2.80 (0.97–8.09) |
| Census-based, women (O’Reilly and Rosato, 2013) | 36,361 | 3.8 | 1.93 (0.44–8.47) |
| **Summary estimate**             | 154,231  | 100    | 1.70 (1.27–2.27) |
| (P = 0%, P = 0.86)               |          |        |   |                       |

a Adjusted for age and sex.

b In census-based study, “own account” employees combined with intermediate SES. Summary hazard ratio in intermediate SES is 1.14 (0.73–1.80) if own account employees were combined with routine occupations and 1.14 (0.73–1.80) if own account employees were excluded.

c In census-based study, “own account” employees combined with intermediate SES. Summary hazard ratio in low SES is 1.47 (1.20–1.81) if own account employees were combined with routine occupations and 1.70 (1.27–2.27) if own account employees were excluded.

The conclusion about “clinical meaningfulness” lacks justification

The WHO/ILO report does not explain why the excess risk of ischaemic heart disease for those working ≥55 h/week was evaluated as “clinically meaningful”, nor did we find a definition for that term in the Navigation Guide or GRADE tools used by the WHO/ILO.

In broad usage, clinical significance refers to a change in a person’s clinical status that is regarded as important, or how much change in disease incidence does the risk factor cause. According to the American Heart Association/American of College Cardiologists Guidelines, (Goff et al., 2014) people with a 10% risk of a major cardiovascular event after 10-years surveillance are denoted as being ‘high risk’. Based on a hazard ratio of 1.13, additional exposure to long working hours would increase their risk by 1.3 percent points. For individuals at ‘intermediate’ (5%) risk, the increase would be only 0.7 percent points, and the change would be even smaller for those with a ‘low’ (2%) risk, 0.3 percent points. These changes in risk due to long working hours are likely to be too modest to justify the description of being ‘clinically meaningful’.

However, small effect sizes can be influential at the population level if a large number of persons are exposed. Population attributable fraction reflects the prevalence of the risk factor in the population and the strength of its association with the outcome being considered. The core assumption in this calculation, however, is that the risk factor has a “causal” link to the outcome. The prevalence of employees working 55 h or more a week in Europe appears relatively low, varying between 1.2% and 16.6% (mean 7.2%) (Kivimäki et al., 2015; Hanmer et al., 2018; Hayashi et al., 2019). Assuming that working fewer than 55 h per week does not increase the risk of ischaemic heart disease, as suggested by the WHO/ILO review, and that the observed association between long working hours and ischaemic heart disease is causal, the report implies but as we dispute, the population attributable fraction for the proportion of heart disease cases avoided by removing all long working hours would be low: between 0.2% and 2.1%, an estimation that is likely to vary markedly by region. For comparison, the corresponding population attributable fraction is 15% to 20% each for established clinically significant cardiovascular disease risk factors, such as hypertension, hypercholesterolaemia and smoking in high-income countries (Yusuf et al., 2020).

In light of these considerations, the WHO/ILO report’s argument that the excess risk of ischaemic heart disease for those working ≥55 h/week is “clinically meaningful” remains insufficiently justified.

5. WHO systematic reviews are authoritative and should be kept to the highest standards

Given its position as the leading global health authority, information disseminated by the WHO can have substantial impact on global health policy. Making far reaching conclusions based on uncertain evidence is risky as it may

- expose people to ineffective strategies that divert attention and resources from preventative approaches that are more effective;
- lead to misleading public health messages that are difficult to correct later; and
- affect confidence in other WHO/ILO guidelines among health professionals, patients and the general public.

The WHO/ILO report on long working hours and ischaemic heart disease is based on a carefully planned registered review protocol, and is well-documented and transparent (Li et al., 2020). However, by drawing definite conclusions, the WHO/ILO expert group has not correctly applied their own framework for assessing the strength of the evidence and they make too far-reaching inferences based on uncertain data. The Navigate Guide requires consistency of findings for a conclusion of “sufficient evidence of harmfulness”. The findings are not compared to standard 35–40 weekly hours) than in the superior-quality studies with low risk of bias for which the estimate was not significantly different from the null (1.08, 95% confidence interval 0.93–1.25).

Given that the overall association between long working hours and ischaemic heart disease was weak and not robust to stratification, multivariable adjustments or exclusion of lower-quality evidence, the findings are uncertain and a conclusion of “sufficient evidence of harmfulness” surely premature.
consistent but differ between subgroups. In addition, bias and confounding cannot be ruled out “with reasonable confidence” when no association is seen after adjustment for covariates and the estimates in studies with low risk of bias are not significantly different from the null. Furthermore, it is not appropriate to provide clinical interpretations without clear supporting data.

Recognising what we do not know will lead to meaningful progress in research and therefore we hope that the interpretations made by the authors of the WHO/ILO report will be discussed more widely in the scientific community. Clearly, the reviewed evidence does not allow evaluation of clinical significance and the findings of socioeconomic interaction suggest that at this stage the conclusion should be restricted to low socioeconomic status occupations only.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgements**

Kivimäki is supported by NordForsk (the Nordic Research Programme on Health and Welfare), the UK Medical Research Council (R024227), the National Institute on Aging (NIA, NIH), US (R01AG056477, R01AG062553), the Academy of Finland (311492) and Helsinki Institute of Life Science, Virtanen by the Academy of Finland (329201, 329226) and the Finnish Work Environment Fund, Nyberg by NordForsk and Batty by the Medical Research Council (MR/P023444/1) and NIA (grant 1R56AG052519-01). There was no direct financial or material support for this commentary.

**References**

Alfredsson, L., Hammar, N., Fransson, E., et al., 2002. Job strain and major risk factors for coronary heart disease among employed males and females in a Swedish study on work, lipids and fibrinogen. Scand. J. Work Environ. Health 28 (4), 238–248.

Bradford, Hill A., 1965. The environment and disease: association or causation? Proc. Royal Soc. Med. 58, 295–300.

De Bacco, D., Pelfrene, E., Clays, E., et al., 2005. Perceived job stress and incidence of coronary events: 3-year follow-up of the Belgian Job Stress Project cohort. Am. J. Epidemiol. 161 (5), 434–441.

Descatha, A., Sembajwe, G., Pega, F., et al., 2020. The effect of exposure to long working hours on ischaemic heart disease: a systematic review and meta-analysis of the WHO/ILO Joint Estimates of the Work-Related Burden of Disease and Injury. Environ. Int. 142, 105739.

Marmot, M.G., Davey Smith, G., Stansfeld, S., et al., 1991. Health inequalities among British civil servants: the Whitehall II study. Lancet 337 (8754), 1387–1393.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 6 (7), e1000097.

Rivara, A.S., Akanbi, M., O’Dwyer, L.C., McHugh, M., 2020. Shift work and long working hours and their association with chronic health conditions: a systematic review of systematic reviews with meta-analyses. PLoS ONE 15 (4), e0231037.

Shea, B.J., Reeves, B.C., Wells, G., et al., 2017. AMSTAR 2: a critical appraisal tool for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 6 (7), e1000097.

Woodruff, T.J., Sutton, P., 2014. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. Environ. Health Perspect. 122 (10), 1007–1014.

Yusuf, S., Joseph, P., Rangarajan, S., et al., 2020. Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. Lancet 395 (10226), 795–808.