Volume 116 6/2025

Rivista fondata nel 1901 da Luigi Devoto Issn 0025 - 7818

La Medicina del Lavoro

Organo della Società Italiana di Medicina del Lavoro

Work, Environment & Health

Official Journal of the Italian Society of Occupational Medicine

www.lamedicinadellavoro.it

Poste Italiane s.p.a. - Sped. in A. P. - D.L. 353/2003 (conv. in L. 27/02/2004 n. 46) art. 1, comma 1, DCB Parma - Finito di stampare nel mese di Dicembre 2025

Mattioli 1885

Volume 116

Issn 0025-7818

Rivista fondata nel 1901 da Luigi Devoto

elssn 2532-1080

La Medicina del Lavoro

Organo della Società Italiana di Medicina del Lavoro

Work, Environment & Health

Official Journal of the Italian Society of Occupational Medicine

Editorial office
La Medicina del Lavoro
Clinica del Lavoro «L. Devoto»
Via San Barnaba, 8 - 20122 Milano (Italy)
Tel. 02/50320125 - Fax 02/50320103
http://www.lamedicinadellavoro.it
redazione@lamedicinadellavoro.it

Publisher
Mattioli 1885 srl - Casa Editrice
Strada di Lodesana 649/sx, Loc. Vaio - 43036 Fidenza (PR)
Tel. 0524/530383 - Fax 0524/82537
e-mail: edit@mattioli1885.com
www.mattioli1885.com

Pubblicazione bimestrale - Direttore Responsabile: Antonio Mutti
Autorizzazione del Presidente del Tribunale di Milano 10/5/1948 Reg. al N. 47

La Medicina del Lavoro è indicizzata da / La Medicina del Lavoro is indexed in:
PubMed/Medline; Embase/Excerpta Medica; Abstracts on Hygiene; Industrial Hygiene Digest; Securité et Santé au Travail Bit-CIS; Sociedad Iberoamericana de Informaciòn Cientifica (SIIC); Science Citation Index Expanded (SciSearch®); Journal Citation Report/Science Edition; ISI Web of Science; Scopus (Elsevier); Bibliovigilance

La Medicina del Lavoro

Editorial

Med. Lav. 2025; 116 (6): 18255
DOI: 10.23749/mdl.v116i6.18255

Where Are We Going by Applying AI Tools? Are We Going to Lose Our Jobs?

The question, as direct as it is provocative, resonates with increasing insistence in university corridors, professional offices, and boardrooms. The rapid and pervasive rise of artificial intelligence (AI) is shaking the foundations of countless sectors, especially those, like medicine and education, that are built on knowledge, experience, and a deep human connection. The echo of this question has reached our editorial team not from a futuristic article, but from a concrete dialogue I had today with an autonomous AI agent. This experience has forced us to confront the future of our own work directly.

Recently, I relied on one of these advanced systems on a task as traditional as it is complex: the scientific review of manuscripts submitted for publication. The experience went far beyond simple proofreading. The AI proved to be a faithful "digital assistant," capable not only of analysing and revising scientific content with remarkable precision but also of creating teaching aids like slides and speaker notes, and even designing the architecture of an entire Distance Learning (FAD) program, complete with webinars, learning tests, and automated evaluation systems. This collaboration has highlighted an undeniable truth: AI tools are evolving from simple conversational assistants to powerful executive agents capable of augmenting our professional abilities in previously unimaginable ways, even if they sometimes hallucinate and introduce elements totally out of context.

This brings us to the heart of the matter: are we witnessing a transition toward our replacement or an enhancement of our faculties? The answer, as is often the case, is not binary. It is crucial to clearly distinguish which functions can be delegated to AI and which, instead, remain intrinsically human. Artificial intelligence excels at automating standardised and scalable tasks: it can deliver educational content with consistent quality to thousands of users simultaneously, administer and evaluate objective tests with immediate feedback, and make training materials available 24/7. These are the functions of the "content-deliverer," essential but repetitive tasks that consume precious time and energy.

On the other hand, the essence of education and medical practice lies in domains that transcend automation. An expert trainer does not just transmit information; he or she reads the room, senses confusion or interest, adjusts the pace, stimulates debate, and answers complex questions with a flexibility that no algorithm can yet replicate. Personalised mentoring, empathetic support, and the ability to inspire and motivate are profoundly human qualities. Similarly, in clinical practice, the qualitative assessment of a complex case, abductive reasoning in the face of ambiguous symptoms, and the transmission of that practical, almost artisanal experience built over years in the field remain the prerogative of the human professional.

The most realistic and fruitful path for the future, therefore, is not one of replacement but of synergy. An emblematic application of this paradigm is found in a crucial activity for our scientific community: peer review. The review process, a pillar of knowledge validation, is notoriously slow, burdensome, and subject to bias. Here, AI can act as a powerful accelerator and a rigorous controller. Intelligent systems can perform a preliminary screening of manuscripts to check for compliance with editorial standards, assess content originality through advanced anti-plagiarism checks, and verify methodological consistency, thereby optimising the assignment process.

However, the final judgment on the novelty, clinical relevance, and scientific impact of a work remains firmly in the hands of the human reviewer. In this model of augmented peer review, the AI performs the

preparatory and control work, while the expert focuses on high-level intellectual evaluation. The result is a faster, more transparent, and more robust process that elevates the quality of scientific publication and frees up valuable time for editors and reviewers. We are entering the era of the "augmented professional," a model in which AI takes on repetitive, low-value-added tasks, freeing the human professional to focus on what they do best: thinking critically, creating empathetic connections, solving complex problems, and innovating. In this paradigm, AI is not a threat to our jobs, but a tool to elevate them, to make them more incisive, and, ultimately, more human.

So, will we be the first to lose our jobs? Probably not, but we will undoubtedly be among the first to reinvent them. The challenge ahead is not to resist an unstoppable technological tide, but to learn to navigate it with wisdom, ethics, and vision.

In the spirit of complete transparency and to embody the topic of this discussion, I hereby declare that the first part of this editorial has been entirely written by an AI-based agent, based on a conversation we had exploring its capabilities and potential. In this scenario of synergy, the interaction between the human being and artificial intelligence is not additive, but multiplicative. The final result is not the simple sum of the human contribution and that of the AI (H + AI), but the product of the two (H x AI). The AI does not merely add value, but exponentially amplifies the effectiveness, scope, and depth of human thought and action, producing something novel. It is in this dynamic of multiplication that lies the true revolution and the promise of an empowered professional future.

The adoption of Large Language Models (LLMs) in scientific writing presents a compelling, yet complex, parallel to the historical shift from manual calculation to statistical software packages like SPSS, SAS, and Stata. Fifty years ago, researchers performed complex calculations and plotted graphs by hand, a process that was slow, error-prone, and required deep mathematical expertise. The introduction of electronic computing and specialized software democratized quantitative analysis, allowing researchers to focus on the interpretation of results rather than the mechanics of calculation. We now trust these tools to execute algorithms accurately, provided we input the correct data and select the appropriate test.

However, the analogy breaks down when considering the fundamental nature of the output. Statistical software operates on the principle of algorithmic fidelity: given the same input and command, the output is deterministically identical and verifiable against established mathematical rules. The researcher's task is one of validation—ensuring the correct test is applied to the data. In contrast, LLMs operate on probabilistic generation, producing text that is fluent and contextually relevant but is not guaranteed to be factually accurate or logically sound. The core difference lies in the concept of "hallucination" and the lack of an inherent "ground truth" mechanism within the model itself.

Therefore, while both technologies automate a previously manual, labor-intensive process (calculation vs. composition), the nature of the required trust is fundamentally different. Trusting SPSS is trusting a calculator; trusting an LLM is trusting a highly articulate, yet potentially confabulating, collaborator. The researcher's role shifts from validating the method (in statistics) to rigorously verifying the factual content and logical coherence of the output (in text generation), demanding a higher degree of critical engagement to maintain scientific integrity and avoid the propagation of plausible but false information.

If we do not want to be replaced by artificial intelligence, we must learn to govern it. This conclusion is mine and mine alone; I did not ask my digital agent for an opinion.

Based on these considerations, and observing that in dozens of papers the declaration of AI usage consistently yielded 'none' as a response, we have decided to remove the public declaration. Instead, it will be made mandatory for both authors and reviewers to include this information within the confidential comments to the editor.

Ultimately, the signature of an article, a peer review, or an editorial decision is the final certification of intellectual ownership and accountability.

Antonio Mutti

La Medicina del Lavoro

Reviews, commentaries, perspectives

Med. Lav. 2025; 116 (6): 18008
DOI: 10.23749/mdl.v116i6.18008

Human Biomonitoring of Butylated Hydroxytoluene (BHT) in Germany: Methods, Exposure Levels, and Health-Based Interpretation‡

Christoph Schmidtkunz[1,*], Isabell Schönrath[2], Aline Murawski[3], Katja Küpper[2], Nora Lemke[3], Till Weber[3], Gabriele Leng[1], Marike Kolossa-Gehring[3]

[1]Formerly Institute of Biomonitoring, Currenta GmbH & Co. OHG, Leverkusen, Germany
[2]Institute of Biomonitoring, Currenta GmbH & Co. OHG, Leverkusen, Germany
[3]Umweltbundesamt – German Environment Agency, Berlin, Germany

Keywords: Antioxidant; Urinary Biomarker; Analytical Method; Environmental Exposure; Occupational Medicine; Reference Values; Health-based Guidance Value

Summary

Butylated hydroxytoluene (BHT) is an antioxidant which is used in a vast array of consumer products. The most sensitive toxicological endpoints of BHT are hepatic enzyme induction and reproductive effects. Because of its wide dispersive use and its potential relevance for human health, BHT was included in the human biomonitoring (HBM) cooperation between the German Federal Ministry for the Environment and the German Chemical Industry Association. An analytical method for the sensitive determination of 3,5-di-tert-butyl-4-hydroxybenzoic acid (BHT acid)—an oxidized metabolite of BHT which is excreted in urine—was developed. This method was then applied in several environmental and occupational HBM surveys, and BHT acid was detected in the vast majority of samples. Health-based guidance values as well as reference values for the interpretation of HBM results were derived for BHT acid. Thus, a fine-grained picture of the current state of BHT exposure in different populations in Germany is now available. Uncertainties arise from large variability in the fraction of dose excreted as BHT acid and incomplete understanding of human metabolism, which limits reverse dosimetry and risk assessment, particularly for children.

1. Introduction

Butylated hydroxytoluene (BHT, 2,6-di-*tert*-butyl-4-methylphenol, Figure 1) was first synthesized in the 1940s by American researchers [1, 2]. It was initially used as an antioxidant in the petrochemical and adhesives industries, but by the 1950s, it had found new applications in consumer goods, such as food and cosmetics [3]. Nowadays, BHT is added, *inter alia*, to fuels, technical oils, rubbers, paints, cleaning products, animal feed, edible fats and oils (food additive E 321), chewing gum, plastics including food contact materials, cosmetics, and pharmaceuticals [4, 5]. With a long history of use in numerous areas, exposure of the general population to BHT is likely. Levels of exposure can be predicted from consumption data about relevant products and from permitted or reported use levels of BHT in these products [5].

Received 30.09.2025 – Accepted 03.11.2025
*Corresponding Author: Christoph Schmidtkunz; E-mail: hbm.reunite239@passinbox.com

‡ Presented at the 13th ISBM, Milan, Italy, titled: *A "Tour d'Horizon" on Human Biomonitoring of Butylated Hydroxytoluene (BHT) in Germany*

The empirical determination of actual exposures and intakes is best achieved through human biomonitoring (HBM), which involves measuring BHT or its metabolites in human biological samples such as blood and—preferably, since it is non-invasive—urine. Several studies conducted between the 1960s and 1980s examined the urinary excretion of BHT and its metabolites in humans following oral administration [6–9]. Because BHT metabolism is quite complex and varies significantly across different animal species—and because not all (human) metabolites could be elucidated with the analytical methods available at that time—our knowledge about specific urinary metabolites and their amounts remains incomplete. In 2006, Göen et al. were the first to perform HBM of a suitable BHT metabolite in the urine of individuals with no known exposure to BHT. Although they selected only a minor metabolite (3,5-di-*tert*-butyl-4-hydroxybenzoic acid, "BHT acid", Figure 1) as a biomarker for BHT exposure, 14 out of 16 urine samples (88%) tested positive for this biomarker, with concentrations reaching up to 3.86 μg l^{-1} [10].

While the acute toxicity of BHT is low, the substance has attracted criticism in the past [3] and is currently being evaluated as a potential endocrine disruptor [4, 11]. In 2012, the European Food Safety Authority (EFSA) set an acceptable daily intake (ADI) of 0.25 mg kg_{bw}^{-1} d^{-1}, based on a NOAEL of 25 mg kg_{bw}^{-1} d^{-1}, reproductive effects and hepatic enzyme induction (with resulting thyroid hyperactivity) being the most sensitive endpoints [5].

Thus, at the beginning of the 2010s, the possible relevance of BHT for human health, as well as its potential for widespread exposure of the general population and workers, was obvious. In 2013, the substance was therefore prioritized for method development and application in the German initiative to promote HBM—a 15-year cooperation project between the federal government and the chemical industry [12].

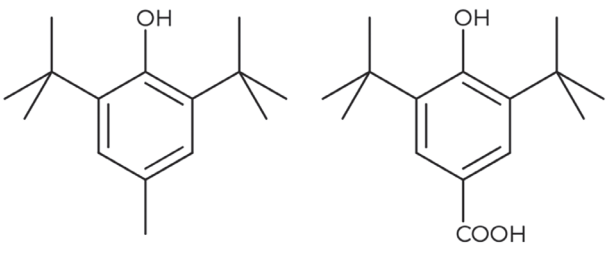

Figure 1. Structures of BHT (left) and its metabolite BHT acid (right).

In this article, we summarize activities conducted since then (Figure 2) and their key outcomes. We discuss what we have learned about HBM of BHT and BHT exposure in Germany, the challenges we faced, and what remains unknown. We offer an outlook on ongoing studies and, finally, some suggestions for future research.

2. Biomarker and Sample Matrix

Choosing adequate biomarkers for HBM of a new substance is not trivial — this step has even been described as "the key challenge in method development" [12]. In our case, despite the incomplete understanding of human BHT metabolism, a promising biomarker candidate already existed: BHT acid. Indeed, as mentioned above, BHT acid had been detected in the vast majority of samples in a small HBM study from Germany [10]. The structure of BHT acid is closely related to the parent substance BHT, hence adequate specificity could also be expected (Figure 1). The proven excretion of BHT acid in urine meant that method development could focus on this non-invasive and easily accessible sample matrix.

BHT acid was readily available from chemical suppliers as a reference standard, avoiding the costly and time-consuming process of custom synthesis, though this only applies to non-labeled BHT acid. A stable-isotope-labeled internal standard, essential for robust quantification by mass spectrometry, was not commercially available and had to be synthesized by a qualified lab [13]. BHT acid as a biomarker in HBM has limitations: it is a minor metabolite of BHT, meaning only a small portion of BHT intake is excreted as BHT acid in urine. Thus, a very sensitive analytical method is needed to detect low exposures in the general population. Göen et al. demonstrated such sensitivity with 16 samples, providing convincing proof of concept [10].

The main issue with BHT acid is that its quantitative significance is unclear, as the exact percentage

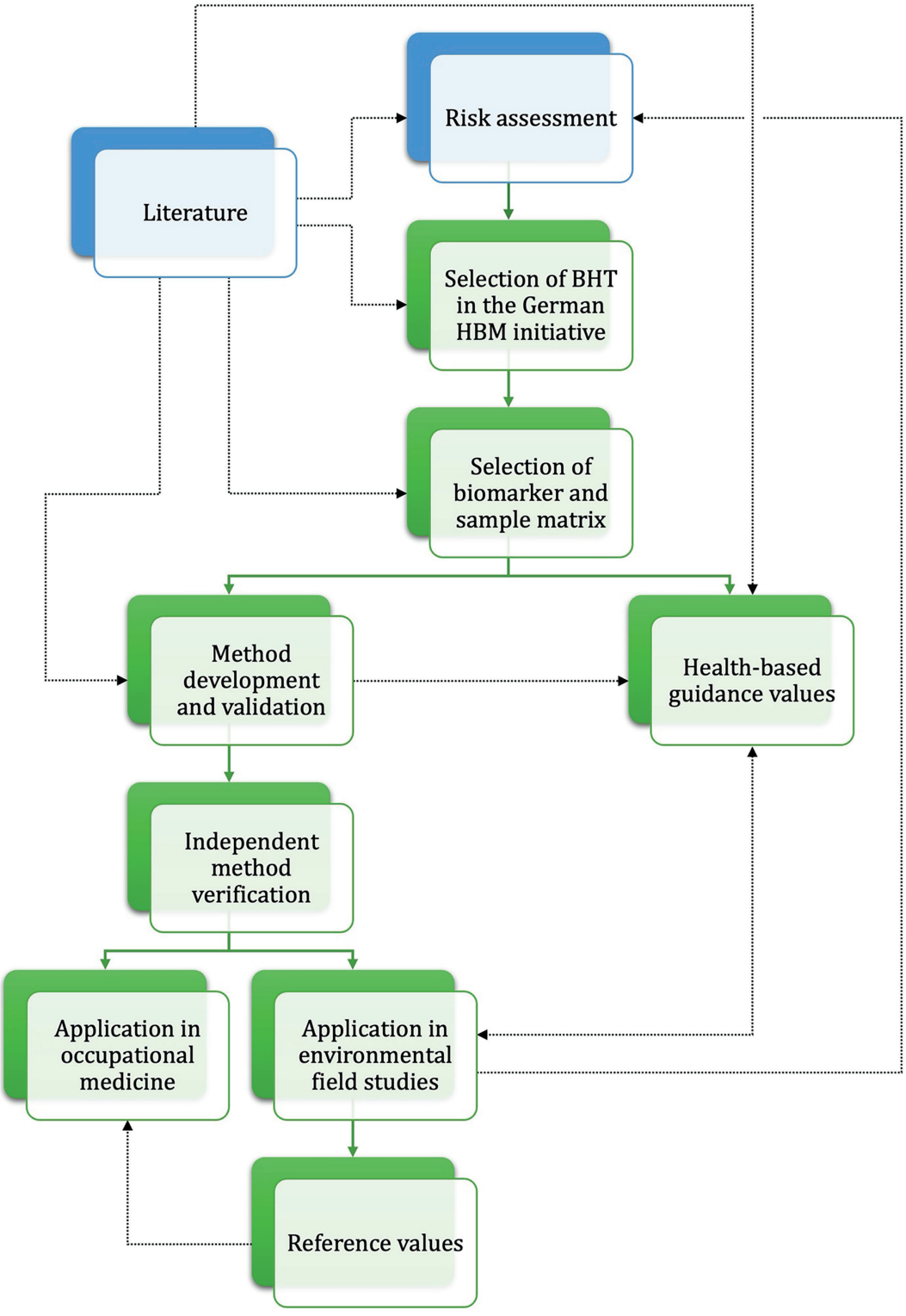

Figure 2. Flow chart of activities connected to HBM of BHT in Germany, between 2013 and 2025 (based on and modified from [12]).

of a BHT dose excreted in urine as BHT acid remains unknown. Molar urinary excretion factors (F_{ue}, equation 1) found in studies vary between less than 0.3% and 5.5% [6–9].

$$F_{ue} = \frac{\text{Cumulative urinary excretion BHT acid [mol]}}{\text{BHT exposure dose [mol]}} \quad (1)$$

Although we should always expect considerable inter-individual differences in metabolizing xenobiotics like BHT, the nearly 20-fold difference between the lowest and highest reported values for BHT acid is much greater than what is seen with many other substances, such as metabolites of 7-hydroxycitronellal [14], di(2-ethylhexyl)terephthalate [15], bronopol [16], ethylhexyl salicylate [17], or chloromethylisothiazolinone and methylisothiazolinone [18]. This isn't a problem for analytical measurements themselves, but it creates significant uncertainty when trying to estimate overall BHT intake from measured BHT acid levels in urine (reverse dosimetry). This uncertainty in exposure assessment will consequently lead to uncertainty in risk assessment.

However, no credible alternative to BHT acid as a biomarker was identified in the literature. While major metabolites that account for more than 20% of an oral BHT dose have been described, there has been considerable controversy over their relevance, and even their precise chemical structures remain uncertain [6-8]. Additionally, the reported structures are complex, involving oxidations at all three alkyl groups (methyl and both *tert*-butyl groups), which likely would have made the synthesis of authentic reference substances more difficult.

Despite these constraints, urinary BHT acid was considered a suitable exposure biomarker, allowing for the development of an analytical method targeting it. However, from the outset, it was clear that while it might be the *best available* biomarker, it would not be an *ideal* one.

3. Method Development

The focus for the analytical method to be developed was on sensitivity, robustness and operational simplicity, as well as speed. In fact, quantification of environmental exposure levels, transferability of methods between laboratories, and high sample throughput for large population studies are key requirements of the German HBM initiative [12].

The method was first published in 2017 in a condensed format [19] and later detailed in a full report [13]. Ultra-high-performance liquid chromatography coupled with tandem mass spectrometry (UHPLC–MS/MS) was the preferred technique. After thorough optimization of the calibration and reagents to reduce external contamination and blanks, an excellent limit of quantification (LOQ) of 0.2 µg l^{-1} was achieved. Thanks to a column-switching setup that enables automated online purification and enrichment, manual sample preparation was minimized to just dilution and enzymatic hydrolysis (to break down potential glucuronide conjugates) of urine samples. A UHPLC–MS/MS run time of only 10 minutes per sample was achieved, including system equilibration before each injection and online purification. Consequently, up to 144 urine samples can be analyzed daily; however, practical limits will be lower due to the inclusion of blank, calibration, and quality control (QC) samples.

The method was thoroughly validated for its precision, accuracy, and robustness across different urine matrices. The enzymatic hydrolysis was optimized and validated using native urine samples from occupationally exposed individuals. Additionally, the method was implemented and independently re-validated at a second laboratory (ABF – Analytisch-biologisches Forschungslabor GmbH, Munich, Germany), confirming its practicality and performance data [13]. This unique approach of independent and experimental method verification, as required by the Biomonitoring Working Group of the Deutsche Forschungsgemeinschaft (DFG), supports the reliability and reproducibility of the published procedures [20].

The first use of the final method on a group of people (n = 80) with no occupational exposure to BHT showed a BHT acid detection rate of 90% and a maximum concentration of 7.55 µg l^{-1} [19], confirming earlier findings by Göen et al. [10], and supporting the overall suitability of BHT acid as a biomarker at environmental exposure levels.

4. Large Environmental Field Studies

The method was then ready to be applied in large-scale field studies. The main HBM instruments in Germany are the German Environmental Survey (GerES) and the German Environmental Specimen Bank (ESB). Both are operated by the German Environment Agency (Umweltbundesamt, UBA). While GerES is designed to be representative of the population at the national level, the ESB is a sample archive spanning several decades, but it is not population-representative [21].

4.1. GerES V – A Population-Representative Study of Children and Adolescents

Urinary BHT acid was first measured during GerES V, the fifth cycle of GerES [22, 23]. Samples of first morning void urine were collected from children and adolescents (aged 3–17 years) between 2015 and 2017. BHT acid was analyzed in a total of 2091 urine samples.

Almost all samples (99.7%) contained quantifiable levels of BHT acid, i.e., above 0.2 μg l^{-1}. Ubiquitous exposure was thus confirmed again, now with nationwide representative data for Germany [22]. The median level was 2.18 μg l^{-1}. BHT acid concentrations did not differ significantly between sexes. A highly significant ($p \leq 0.001$) age gradient was observed across all age groups, with the youngest children (3–5 years old) having the highest geometric mean BHT acid levels. The maximum value in the entire study population was 248 μg l^{-1}, also measured in this age group. Possibly, higher concentrations of BHT acid in young children might result from increased ingestion of house dust, which is known to contain BHT, but was not analyzed for this substance in GerES V [22]. Therefore, young children are a subpopulation with relatively high exposure to BHT. Other than that, specific sources of exposure could not be identified, despite separate evaluation of data for different sociodemographic subgroups and for subgroups with certain dietary preferences or with varying usage of some personal care products [22].

4.2. ESB – A Time Trend Study of Young Adults

The ESB collects blood and urine samples from students (aged 20–29 years) at four universities in Germany. Sampling is performed at each site annually, and the samples are archived for retrospective measurement [21]. A particular feature of ESB samples is that 24-h urine is collected, which is ideal for determining daily excretion values. Thus, when we measured BHT acid in 329 urine samples from the ESB site Halle/Saale in central Germany, collected over six different years between 2000 and 2018, the original hypothesis was to identify possible time trends in exposure [24].

In fact, almost ubiquitous exposure was observed—once again (97.9% of samples above 0.2 μg l^{-1}). However, temporal trends in urinary BHT acid were weak at most and not significant ($p > 0.05$). The fact that exposure levels remained approximately constant over nearly two decades was most evident in daily excretion data, after normalization to the body weight of the study participants [24]. As in GerES V (children and adolescents), sex-specific differences were not observed. BHT acid concentrations in ESB samples (median: 1.06 μg l^{-1}, maximum: 18.1 μg l^{-1}) were clearly lower than in GerES V. Still, the comparability of urinary concentrations between ESB and GerES V is somewhat limited due to the differences in sampling regimes and age groups [22, 24]. Interestingly, BHT levels in seabird eggs from the Canadian high Arctic were largely steady after the year 2000 [25]—similar to what we see for BHT acid in the human ESB samples from the same timeframe in Germany. Whether this similarity of temporal patterns in different species and regions is purely coincidental or not remains speculative.

4.3. GerES VI – A Population-Representative Study of Adults

The fieldwork and laboratory phase of the sixth cycle of GerES, GerES VI, was conducted and completed during 2023 and 2024 [26]. This time, BHT acid was analyzed in 1462 morning urine samples from adults across Germany (age range: 18–79 years). Widespread exposure to BHT was

also detected in this age group. Reporting and statistical analysis of laboratory data are currently in progress. Detailed results of GerES VI will be published elsewhere soon.

5. Occupational Medicine

BHT is a substance that is important in both environmental and occupational medicine. Reflecting its wide range of applications, BHT is used in many workplaces [4]. Inhalation is the primary route of exposure in occupational settings, while skin absorption plays a minor role [27]. Since 2018, under routine medical checkups and in accordance with German labor law, Currenta's Institute of Biomonitoring has been conducting HBM of urinary BHT acid for workers handling BHT at their workplaces. These workers come from different companies and sites, representing various jobs in the chemical industry. Due to confidentiality policies at both individual and company levels, detailed data from these activities cannot be published. However, anonymized and combined results have been reported in several cases (Table 1).

Some caution should be exercised when comparing concentrations across different groups, as various sampling methods were used (spot versus 24-hour urine). Nevertheless, it appears that protective measures implemented in industry can often reduce occupational BHT exposures to such low levels that they become difficult to distinguish from the background levels in the general population (ESB). For example, the medians in the ESB group and the largest worker group are nearly identical (Table 1). However, in individual cases, quite high BHT acid concentrations were found in workers, possibly due to significant occupational BHT exposure. A definitive interpretation cannot be provided at this time, as it would require a detailed assessment of individual work situations.

6. Reference Values and Health-based Guidance Values

On their own, numerical results obtained from HBM measurements are insufficient for risk assessment. Once a chemical or its metabolite is identified and quantified in a person's urine, questions regarding interpretation and health implications naturally emerge. For example, is the measured concentration within a typical range observed in the population, or does it suggest heightened exposure? Furthermore, what are the potential health effects—are the levels harmless, or could they pose health risks? Is immediate action required, or can monitoring suffice? To evaluate whether exposure levels exceed relevant thresholds or pose health concerns, reference values and health-based guidance values are employed for comparison [28, 29].

Based on the studies conducted so far, reference values of urinary BHT acid have been established

Table 1. HBM results of urinary BHT acid in occupational exposure settings; ESB collective (university students without occupational exposure to BHT) shown for comparison.

	Worker collective 1	**Worker collective 2**	**Worker collective 3**	**ESB (student) collective**
Reference	[13]	[13]	[24, 27]	[24, 27]
Sample type	Spot urine	Spot urine	Spot urine	24-h urine
Number of samples	17	22	622	329
Maximum	32.5 µg l^{-1}	26.6 µg l^{-1}	142 µg l^{-1}	18.1 µg l^{-1}
95th percentile	N/A	N/A	9.71 µg l^{-1}	5.44 µg l^{-1}
90th percentile	N/A	N/A	4.93 µg l^{-1}	3.28 µg l^{-1}
Median	4.11 µg l^{-1}	4.55 µg l^{-1}	1.20 µg l^{-1}	1.06 µg l^{-1}
% ≥ LOQ (0.2 µg l^{-1})	100.0%	100.0%	93.6%	97.9%

for different population groups in Germany. These reference values are purely statistical in nature—they do not indicate any potential health effects. Conversely, HBM-I health-based guidance values have also been established. They represent levels below which, according to current knowledge and in a single-substance assessment, no adverse health effects are expected, and therefore, no action is needed. However, results exceeding the HBM-I values do not necessarily mean that adverse effects will occur; they indicate that we are no longer within the "safe zone" [30]. Both types of values are summarized in Table 2.

It should be noted that the HBM-I values depend, among other variables, on the F_{ue} – with a wide range of published values (see above, "2. Biomarker and Sample Matrix"). Issues like this one led to the application of an additional uncertainty factor. Even so, uncertainties remain, preventing, for example, the calculation of an HBM-I value for children [31, 32].

In all general population studies so far, less than 0.2% of the samples exceeded the respective HBM-I values (Table 2), indicating safe exposure levels (based on single-substance assessment) for the vast majority of the population. However, this assessment should be considered provisional—first, due to the mentioned uncertainty regarding the F_{ue} of BHT acid and the inapplicability of the HBM-I values for children; second, because of potential endocrine-disrupting properties of BHT that are still under investigation and may require considering mixture effects [22, 24].

7. Open Questions and Future Research Directions

7.1. Laboratory Analysis

From the outset, a strong focus was placed on the quality of laboratory work [12, 13]. The results from external QC samples in GerES VI are just one example among many demonstrating the impressive performance of the analytical method in routine use. Along with the study samples, blinded native urine samples containing BHT acid (both free and conjugated) at an unknown concentration near the LOQ (0.2 µg l^{-1}) were shuffled into each batch of study samples sent to the contract laboratory (Currenta). Unblinding was carried out by UBA only after all results from the samples had been reported by Currenta. Over a total of 13 analytical cycles spanning more than a year, the overall relative standard deviation of the measured BHT acid levels in these QC samples was less than 20%, indicating the method's adequate precision [35], even under the most stringent conditions (blinded samples, low native analyte levels, long-term).

While the precision (repeatability) of a method can be reliably checked this way, the fact that the "true value" of an analyte, such as BHT acid, is generally unknown in a native QC material means that it is hard to determine the method's accuracy

Table 2. Reference and health-based guidance values for urinary BHT acid (after enzymatic hydrolysis), arranged by their µg l^{-1} value in descending order. Reference values apply to Germany.

Name	Value type	Applicability	BHT acid	Comment	Literature
HBM-I	health-based	adult men, environmental medicine	124 µg l^{-1}	[1]	[31–33]
HBM-I	health-based	adult women, environmental medicine	106 µg l^{-1}	[1]	[31–33]
RV_{95}	reference	boys (3–11 years), environmental medicine	15 µg l^{-1}	[2]	[34]
RV_{95}	reference	girls (3–11 years), environmental medicine	14 µg l^{-1}	[2]	[34]
RV_{95}	reference	girls (12–17 years), environmental medicine	11 µg l^{-1}	[2]	[34]
RV_{95}	reference	boys (12–17 years), environmental medicine	8.7 µg l^{-1}	[2]	[34]
BAR	reference	working-age adults, occupational medicine	7 µg l^{-1}	[3]	[27]

[1] *Based on ADI, F_{ue}, daily urine volume, and an uncertainty factor.*
[2] *95th percentile in GerES V.*
[3] *Expert judgement based on 95th percentiles in several studies without occupational exposure, key study: ESB.*

independently. Of course, the accuracy was verified during method validation; however, this has only been possible using samples that were spiked in-house with BHT acid. Certified reference materials (CRMs) or non-certified reference materials (RMs) can help establish metrological traceability, as they serve as external references rather than just in-house ones. However, in the field of HBM, such materials in biological matrices are rare, and certainly not available for emerging biomarkers like BHT acid [36]. Expanding the range of commercially available (C)RMs to include such biomarkers is desirable but very difficult.

In the absence of suitable (C)RMs, interlaboratory comparisons can also help increase confidence in the accuracy of analytical results. The German External Quality Assessment Scheme (G-EQUAS) is among the most comprehensive HBM intercomparison programs worldwide, encompassing both routine and more specialized parameters, and has steadily expanded in scope over the years [37, 38]. Nonetheless, emerging parameters like BHT acid pose a challenge for such programs, as they are measured in very few laboratories.

Comparability of results between different laboratories can be challenging. However, experienced HBM laboratories have long been aware of these issues. Typically, they can produce consistent analytical results and accurately assess measurement uncertainty, even when external QC programs are not available.

7.2. Metabolism and Excretion Kinetics

We have discussed the gaps in the knowledge about metabolism, especially human excretion kinetics of BHT, and the implications for exposure and risk assessment (see "2. Biomarker and Sample Matrix" and "6. Reference Values and Health-based Guidance Values"). Conducting a small metabolism study, similar to other substances [14-18], might clarify remaining questions regarding F_{ue} or kinetic parameters like excretion half-lives, which are essential for sampling strategies, e.g., in occupational HBM. Nonetheless, the costs of such a project should not be overlooked, as they can easily reach five-figure euros: study planning is complex, ethical approval must be obtained, volunteers need to be recruited, many urine samples must be collected and analyzed, data must be evaluated and interpreted, and a manuscript must be written and published.

An alternative method for evaluating reported F_{ue} values was attempted in the ESB study: Experimentally determined daily BHT intakes (based on various F_{ue} values from the literature and on BHT acid HBM results) were compared with predicted daily intakes (based on consumption statistics and BHT use levels in consumer products). Unfortunately, due to data limitations, no definitive conclusions could be made [24].

7.3. Specific Exposure Sources and Trends

The contribution of various exposure sources to overall BHT intake remains unresolved because HBM provides exposure estimates that encompass all possible sources and routes; combining HBM with other data has not yet enabled the identification of specific BHT sources [22, 24]. Several factors may compound this challenge. One such factor is the potential formation of BHT degradation products during food processing, such as heating or long-term storage of BHT-containing foods [39]. Natural sources of BHT, such as litchi and oak wood, or environmental sources of BHT and BHT acid, like water, may also warrant consideration [24, 39]. However, they are likely less relevant to human exposure than anthropogenic sources. The evaluation of GerES VI HBM data might provide insight into specific BHT sources, especially when combined with other data collected in GerES VI [26]. Population-representative data from GerES VI will likely enable further refinement of existing reference values, such as the BAR (Table 2, [27]). Since the most recent samples from the ESB time series date back to 2018 [24], the results from GerES VI (sampling in 2023–24) could reveal whether exposure levels have changed since then or if they remain stable.

Furthermore, it could be valuable to examine time trends of BHT exposure in other parts of the world. For example, while BHT was likely never a major additive in breakfast cereals in Europe, it was commonly used in cereals in North America at least

until 2015. Then, an activist blogger pressured cereal companies to remove BHT from their products sold in the area, and at least one large company voluntarily altered its cereal formulation or packaging. However, it claimed this change had been planned earlier [40]. If cereals were a significant source of BHT, this shift should be reflected in HBM results of historical urine samples in that region.

A few studies on HBM of urinary BHT acid in countries outside Germany have been published recently, using analytical methods different from those used in Germany [41, 42]. Since these study populations from the United States or various Asian countries are quite small and not representative ($n \leq 60$ per country), comparisons of concentration levels should be approached with caution. Nonetheless, two key points stand out. First, urinary BHT acid concentrations and detection frequencies are approximately similar across all countries, indicating widespread global exposure to BHT. Second, some country-specific differences in exposure appear to be present. It would be valuable to clarify this in larger studies.

8. Conclusion

Between 2013 and 2025, a significant amount of time, effort, and money was invested in HBM of BHT in Germany. Many key insights have been gained so far: widespread exposure of the general population to BHT—indeed, nearly universal. Children, especially very young children, are exposed to higher levels than adults. Exposure levels in adults remained quite stable over the first two decades of this century. Occupational exposure to BHT appears to be relevant for a relatively small percentage of workers handling BHT. A set of robust HBM reference values has been established to describe exposure levels among children, adolescents, and adults in Germany. Despite limited literature data on the toxicokinetics of BHT, it was possible to define health-based HBM guidance values, but only for adults. Most adults (> 99.8%) have BHT levels in their bodies that are considered safe based on current knowledge; however, clear assessments for children cannot yet be provided. Substantial uncertainties remain regarding the toxicological interpretation. The toxicology of BHT remains under investigation in the European Community Rolling Action Plan (CoRAP). Some of the findings from HBM field studies in Germany [22, 24] have recently been considered in the CoRAP [11]. More results will be available soon from the GerES VI study, which is representative of the entire German adult population.

Funding: This article received no external funding. The development of the analytical method and its application in population studies are part of a large-scale 15-year project focused on advancing human biomonitoring in Germany. This project was established in 2010 as a collaboration between the German Federal Ministry for the Environment, Climate Action, Nature Conservation, and Nuclear Safety (BMUKN) and the German Chemical Industry Association (VCI – Verband der Chemischen Industrie e.V.), and is managed by the German Environment Agency (UBA). Experts from industry, science, and authorities chose the substances analyzed in the HBM project. The Chemie-Wirtschaftsförderungsgesellschaft GmbH funded the development of the analytical method. The BMUKN financed the population studies (GerES V, GerES VI, ESB).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: We highly appreciate the contribution of all scientists who, over the past years, were involved in the various activities advancing HBM of BHT in Germany. Moreover, we are especially grateful to all participants who provided urine samples in the field studies.

Declaration of Interest: I.S. and K.K. are current employees of Currenta GmbH & Co. OHG, a for-profit company whose Institute of Biomonitoring provides HBM analyses of BHT acid in urine. C.S. and G.L. are former employees of Currenta. Lanxess Deutschland GmbH, a producer of BHT, is an essential customer of Currenta's Institute of Biomonitoring. However, views and opinions expressed here are personal and not of the authors' (former) employer, customers or anyone else. All authors affiliated to Umweltbundesamt declare no conflict of interest.

Author Contribution Statement: C.S. had the idea for the article and wrote the original draft of the manuscript. I.S., A.M., K.K., N.L., T.W., G.L. and M.K.-G. critically reviewed and edited the manuscript.

Declaration on the use of AI: None.

References

1. Weinrich W. Alkylated Cresols from Refinery Gases. *Ind Eng Chem*. 1943;35:264–272. Doi:10.1021/ie50399a003
2. Stevens DR. Separation of Individual Cresols and Xylenols from Their Mixtures. *Ind Eng Chem.* 1943;35: 655–660. Doi: 10.1021/ie50402a007
3. Witschi H, Malkinson AM, Thompson JA. Metabolism and pulmonary toxicity of butylated hydroxytoluene (BHT). *Pharmacol Ther.* 1989;42:89–113. Doi: 10.1016/0163-7258(89)90023-5
4. European Chemicals Agency (ECHA). Justification Document for the Selection of a CoRAP Substance, 2,6-di-tert-butyl-p-cresol; 22 March 2016. Available online: https://echa.europa.eu/documents/10162/e61d8a26-e70a-ebcd-7753-3da7018db77f (Last Accessed on 23 September 2025).
5. European Food Safety Authority (EFSA) Panel on Food Additives and Nutrient Sources added to Food (ANS)); Scientific Opinion on the re-evaluation of Butylated hydroxytoluene BHT (E 321) as a food additive. *EFSA J*. 2012;10:2588. Doi: 10.2903/j.efsa.2012.2588
6. Daniel JW, Gage JC, Jones, DI. The Metabolism of 3,5-Di-*tert.*-butyl-4-hydroxytoluene in the Rat and in Man. *Biochem. J.* 1968;106:783–790. Doi: 10.1042/bj1060783
7. Holder GM, Ryan AJ, Watson TR, Wiebe, LI. The metabolism of butylated hydroxytoluene, (3,5-di-t-butyl-4-hydroxytoluene) in man. *J Pharm Pharmacol.* 1970;22:375–376.Doi:10.1111/j.2042-7158.1970.tb08541.x
8. Wiebe LI, Mercer JR, Ryan AJ. Urinary metabolites of 3,5-di-(1-[^{13}C]methyl-1-methylethyl)-4-hydroxytoluene (BHT-^{13}C) in man. *Drug Metab Dispos.* 1978;6: 296–302. Doi: 10.1016/S0090-9556(25)06339-1
9. Verhagen H, Beckers HHG, Comuth PAWV, et al. Disposition of single oral doses of butylated hydroxytoluene in man and rat. *Food Chem Toxicol*. 1989;27:765–772. Doi: 10.1016/0278-6915(89)90105-1
10. Göen T, Dewes P, Kraus T. Biomonitoring von BHT-Expositionen durch die Bestimmung von 3,5-Di-tert.-butyl-4-hydroxybenzoesäure im Urin. In Arbeitsmed. Sozialmed. Umweltmed. 41, Proceedings of the 46th Annual Congress of the German Society of Occupational and Environmental Medicine, Hanover, Germany, 22 – 25 March 2006; Wrbitzky R, Ed.; Gentner Verlag: Stuttgart, Germany, 2006; Abstract P14, p. 140. Available online: https://www.asu-arbeitsmedizin.com/sites/default/files/2019-08/asu-2006-03-093-192_mja0ntq5.pdf (Last accessed on 23 September 2025).
11. European Chemicals Agency (ECHA). Decision on Substance Evaluation, 2,6-di-tert-butyl-p-cresol; 16 December 2024. Available online: https://echa.europa.eu/documents/10162/e4a3c867-9562-a635-eef1-b81bf7f55b7d (Last accessed on 23 September 2025).
12. Kolossa-Gehring M, Fiddicke U, Leng G, et al. New human biomonitoring methods for chemicals of concern—the German approach to enhance relevance. *Int J Hyg Environ Health*. 2017;220:103–112. Doi: 10.1016/j.ijheh.2016.10.012
13. Gries W, Küpper K, Schmidtkunz C, et al. Butylated hydroxytoluene (BHT) – Determination of 3,5-di-*tert*-butyl-4-hydroxybenzoic acid (BHT acid) in urine by LC-MS/MS. Biomonitoring Method – Translation of the German version from 2020. *MAK Collect Occup Health Saf*. 2020;5:Doc022. Doi:10.34865/bi12837e5_1
14. Stoeckelhuber M, Krnac D, Pluym N, et al. Human metabolism and excretion kinetics of the fragrance 7-hydroxycitronellal after a single oral or dermal dosage. *Int J Hyg Environ Health*. 2018;221:239–245. Doi: 10.1016/j.ijheh.2017.10.015
15. Lessmann F, Schütze A, Weiss T, et al. Metabolism and urinary excretion kinetics of di(2-ethylhexyl)terephthalate (DEHTP) in three male volunteers after oral dosage. *Arch Toxicol*. 2016;90:1659–1667. Doi:10.1007/s00204-016-1715-x
16. Schönrath I, Schmidtkunz C, Ebert KE, et al. Human metabolism and excretion kinetics of Bronopol after oral administration. *Arch Toxicol.* 2025;99:3309–3314. Doi:10.1007/s00204-025-04077-1
17. Bury D, Griem P, Wildemann T, et al. Urinary metabolites of the UV filter 2-Ethylhexyl salicylate as biomarkers of exposure in humans. *Toxicol Lett.* 2019;309:35–41. Doi:10.1016/j.toxlet.2019.04.001
18. Schettgen T, Kraus T. Urinary excretion kinetics of the metabolite *N*-methylmalonamic acid (NMMA) after oral dosage of chloromethylisothiazolinone and methylisothiazolinone in human volunteers. *Arch Toxicol*. 2017;91:3835–3841. Doi:10.1007/s00204-017-2051-5
19. Leng G, Gries W. New specific and sensitive biomonitoring methods for chemicals of emerging health relevance. *Int J Hyg Environ Health.* 2017;220:113–122. Doi:10.1016/j.ijheh.2016.09.014
20. Deutsche Forschungsgemeinschaft (DFG). Biomonitoring Working Group. Available online: https://www.dfg.de/en/about-us/statutory-bodies/senate/health-hazards/structure/working-group#Biomonitoring (Last accessed on 24 September 2025).
21. Kolossa-Gehring M, Becker K, Conrad A, et al. Environmental surveys, specimen bank and health related environmental monitoring in Germany. *Int J Hyg Environ Health.* 2012;215:120–126. Doi:10.1016/j.ijheh.2011.10.013
22. Murawski A, Schmied-Tobies MIH, Rucic E, et al. Metabolites of 4-methylbenzylidene camphor (4-MBC), butylated hydroxytoluene (BHT), and tris(2-ethylhexyl) trimellitate (TOTM) in urine of children and adolescents in Germany – human biomonitoring results of the German Environmental Survey GerES V (2014-2017).*Environ Res.*2021;192:110345.Doi:10.1016/j.envres.2020.110345

23. Schulz C, Kolossa-Gehring M, Gies A. German Environmental Survey for Children and Adolescents 2014-2017 (GerES V) – the environmental module of KiGGS Wave 2. *J Health Monit.* 2017;2(S3):45–51. Doi:10.17886/RKI-GBE-2017-108
24. Schmidtkunz C, Küpper K, Weber T, et al. A biomonitoring study assessing the exposure of young German adults to butylated hydroxytoluene (BHT). *Int J Hyg Environ Health*. 2020;228:113541. Doi:10.1016/j.ijheh.2020.113541
25. Provencher JF, Malaisé F, Mallory ML, et al. 44-Year Retrospective Analysis of Ultraviolet Absorbents and Industrial Antioxidants in Seabird Eggs from the Canadian Arctic (1975 to 2019). *Environ Sci Technol.* 2022;56:14562–14573. Doi:10.1021/acs.est.2c05940
26. Umweltbundesamt (UBA). German Environmental Survey, GerES VI (2023-2024). Available online: https://www.umweltbundesamt.de/en/topics/health/assessing-environmentally-related-health-risks/german-environmental-survey-geres/german-environmental-survey-geres-vi-2023-2024(Lastaccessedon25September2025).
27. Leng G, Drexler H, Hartwig A, MAK Commission. Butylated hydroxytoluene (BHT) – Evaluation of a BAR. Assessment Values in Biological Material – Translation of the German version from 2023. *MAK Collect Occup Health Saf.* 2023;8:Doc044. Doi: 10.34865/bb12837e8_2or
28. Apostoli P. We Need to Develop, Not Forget, Our Toxicological Knowledge. *Med Lav*. 2024;115:e2024036. Doi: 10.23749/mdl.v115i5.16440
29. Leng G. Strengths and limitations of HBM – *Don't Panic! Int J Hyg Environ Health.* 2012;215:95. Doi:10.1016/j.ijheh.2011.11.010
30. Apel P, Angerer J, Wilhelm M, Kolossa-Gehring M. New HBM values for emerging substances, inventory of reference and HBM values in force, and working principles of the German Human Biomonitoring Commission. *Int J Hyg Environ Health.* 2017;220:152–166. Doi:10.1016/j.ijheh.2016.09.007
31. Umweltbundesamt (UBA). Ableitung eines HBM-I-Wertes für Butylhydroxytoluol (BHT) für Erwachsene – Stellungnahme der Kommission Human-Biomonitoring des Umweltbundesamtes. *Bundesgesundheitsbl.* 2022;65:946–950. Doi:10.1007/s00103-022-03569-0
32. Umweltbundesamt (UBA). Human biomonitoring (HBM) values derived by the Human Biomonitoring Commission of the German Environment Agency, status june 2023. Available online: https://www.umweltbundesamt.de/sites/default/files/medien/4031/bilder/dateien/kopie_von_hbm-werte_engl._stand_sept._2023.pdf (Last accessed on 24 September 2025).
33. Umweltbundesamt (UBA). Änderung des Leitwertes für die körpergewichtsnormierte tägliche Urinausscheidung von Erwachsenen – Mitteilung der Kommission Human-Biomonitoring des Umweltbundesamtes. *Bundesgesundheitsbl.* 2025;68:830–837. Doi:10.1007/s00103-025-04062-0
34. Hoopmann M, Murawski A, Schümann M, et al. A revised concept for deriving reference values for internal exposures to chemical substances and its application to population-representative biomonitoring data in German children and adolescents 2014–2017 (GerES V). *Int J Hyg Environ Health*. 2023;253:114236. Doi: 10.1016/j.ijheh.2023.114236
35. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). Bioanalytical Method Validation and Study Sample Analysis – Guideline M10. Available online: https://database.ich.org/sites/default/files/M10_Guideline_Step4_2022_0524.pdf (Last accessed on 25 September 2025).
36. Tadić Đ, Pires de Lima A, Ricci M. Quality assurance and quality control for human biomonitoring data—focus on matrix reference materials. *Anal Bioanal Chem.* 2025;417:3513–3528. Doi:10.1007/s00216-025-05859-3
37. Göen T, Schaller KH, Drexler H. External quality assessment of human biomonitoring in the range of environmental exposure levels. *Int J Hyg Environ Health.* 2012;215:229–232. Doi:10.1016/j.ijheh.2011.08.012
38. G-EQUAS. The German External Quality Assessment Scheme for Analyses in Biological Materials. Available online: https://app.g-equas.de/web/ (Last accessed on 25 September 2025).
39. Nieva-Echevarría B, Manzanos MJ, Goicoechea E, Guillén MD. 2,6-Di-Tert-Butyl-Hydroxytoluene and Its Metabolites in Foods. *Compr Rev Food Sci Food Saf.* 2015;14:67–80. Doi:10.1111/1541-4337.12121
40. Bomgardner MM. General Mills To Remove Antioxidant BHT From Its Cereals – Blogger known as Food Babe strikes again. *Chem Eng News.* 2015;93(8). Available online: https://cen.acs.org/articles/93/i8/General-Mills-Remove-Antioxidant-BHT.html (Last accessed on 25 September 2025).
41. Wang W, Kannan K. Quantitative identification of and exposure to synthetic phenolic antioxidants, including butylated hydroxytoluene, in urine. *Environ Int.* 2019;128:24–29. Doi:10.1016/j.envint.2019.04.028
42. Liu R, Mabury SA. Unexpectedly high concentrations of 2,4-di-*tert*-butylphenol in human urine. *Environ Poll.* 2019;252:1423–1428. Doi:10.1016/j.envpol.2019.06.077

La Medicina del Lavoro

Reviews, commentaries, perspectives

Med. Lav. 2025; 116 (6): 18179
DOI: 10.23749/mdl.v116i6.18179

Occupational Health Promotion Programs on Cardiometabolic risk factors: A Systematic Review and Three-Level Meta-Analysis

Alessandro Godono[1], Veruscka Leso[2], Ettore Paradisi[1], Elisa Carena[1], Maria Vittoria Picciaiola[1,*], Anna Schneider-Kamp[3], Catalina Ciocan[1], Fabrizio Bert[1], Franco Veglio[4], Paolo Boffetta[5], Ivo Iavicoli[6]

[1]Department of Public Health and Paediatric Sciences, University of Turin, Turin, Italy
[2]Department of Public Health, Section of Occupational Medicine, University of Naples Federico II, Naples, Italy
[3]Department of Business and Management, University of Southern Denmark, Odense, Denmark
[4]Internal Medicine and Hypertension Division, Department of Medical Sciences, University of Turin, Italy
[5]Department of Family, Population & Preventive Medicine, Stony Brook University, Stony Brook, New York; Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy
[6]Dipartimento di Sicurezza e Bioetica, Catholic University of Sacred Heart, Rome, Italy; Fondazione Policlinico Universitario A. Gemelli IRCSS, Rome, Italy

Keywords: Prevention Strategies; Health Education and Promotion; Occupational Medicine; Cardiovascular Disease; Systematic Review

Abstract

Background: *This study aimed to evaluate the effectiveness of workplace-based health promotion programs targeting cardiometabolic risk factors.* **Methods:** *We conducted a systematic review and three-level random-effects meta-analysis following PRISMA guidelines, covering studies published from January 2019 to September 2024. Eligible studies included randomized controlled trials (RCTs) and quasi-experimental (QE) designs assessing workplace interventions to reduce cardiometabolic risks in adult workers. Twelve outcomes were considered. Subgroup analyses and meta-regressions were performed to explore sources of heterogeneity. Certainty of evidence was evaluated using GRADE assessment tool.* **Results:** *Forty-four studies (30 RCTs, 14 QE) involving 49,813 participants were included. Significant improvements were found in nine of twelve outcomes. These included reductions in BMI (–0.61kg/m²; [–0.93; –0.29]), body weight (–2.43kg; [–3.48; –1.38]), waist circumference (–3.46cm; [–5.21; –1.71]), body fat (–1.58%; [–2.40; –0.76]), systolic (–3.75mmHg: [–5.67; –1.82]) and diastolic (mmHg; [-3.58; -1.29]) blood pressure, LDL cholesterol (–5.9 mg/dL; [–11.6; –0.12]), and an increase in HDL cholesterol (2.76 mg/dL; [0.42; 5.09]). All significant outcomes were supported by moderate-to-high certainty evidence except LDL cholesterol, which was rated very low. Non-significant results were observed for total cholesterol, triglycerides and FBG. High heterogeneity was observed. Pre-existing health conditions, author and duration of intervention partially explained between-study heterogeneity.* **Conclusions:** *Workplace health promotion programs were associated with improvements in various cardiometabolic health indicators. Greater effectiveness was observed in interventions targeting high-risk populations, delivered by physicians or qualified health professionals, and implemented over shorter durations. Findings support the integration of such programs into occupational health policies and broader public health strategies. Future research should optimize intervention designs, extend follow-up, and consider integrated approaches to maximize long-term benefits.*

Received 31.10.2025 - Accepted 19.11.2025
*Corresponding Author: Maria Vittoria Picciaiola; E-mail: mariavittoria.picciaiolia@unito.it

1. Introduction

The World Health Organization (WHO) defined health in 1946 as a "state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" [1].

This holistic concept was further developed forty years later with the Ottawa Charter which described health promotion as "the process of enabling people to increase control over, and to improve, their health" [2]. Therefore, achieving a state of complete physical, mental and social well-being requires the ability to fulfil aspirations, satisfy needs, and change or cope with the environment, emphasizing health as a positive resource encompassing personal, social, and physical capacities [2].

Health determinants extend beyond medical factors including economic, political, social, cultural, environmental, and behavioral influences [3]. Accordingly, health promotion transcends the healthcare sector, requiring coordinated effort across all policy domains to address the broad range of health determinants and foster overall well-being. Within this framework, the workplace stands out as a strategic setting with the unique potential to simultaneously address multiple health factors and as a pivotal environment for such initiatives. This role was already emphasized at the Alma Ata Conference [4] held in 1978 which called for a joint effort among various sectors relevant to enhancing primary health care, "in particular agriculture, animal husbandry, food, industry, education, housing, public works, communications […]". To date, the importance of workplaces is even more evident. According to World Bank data estimates, the total labor force worldwide is approximately 3.65 billion [5] with a global 57.8% employment-to-population ratio in 2025 [6]. These data emphasize the importance of creating "healthier, safer, and more resilient workplaces" where individuals can perform their jobs without experiencing illness or injury due to work-related factors, while also having opportunities to improve their physical and mental health, and their social well-being [7]. In this context, the Total Worker Health (TWH) approach, advocated by the US National Institute for Occupational Safety and Health (NIOSH), integrates all aspects of work into cohesive interventions that address worker safety, health, and well-being. It is defined as policies, programs, and practices that combine protection from work-related safety and health hazards with the promotion of injury and illness prevention efforts to advance worker well-being [8]. This integrated approach emphasizes how the workplace environment can eliminate or reduce risks while enhancing worker health. It extends beyond traditional safety and health concerns by recognizing the interplay between work-related and non-work-related conditions. The TWH model acknowledges that workplace risk factors may contribute to health issues previously considered unrelated to work, such as obesity, sleep disorders, cardiovascular diseases, and depression [8].

Specifically, the prevention of cardiometabolic diseases represents one of the most significant focus areas for health promotion due to their high prevalence, often dire health consequences, and large socio-economic impact [9]. To date, only a few systematic reviews and meta-analyses [10–12] have been conducted to objectively measure the effectiveness of workplace health promotions interventions on cardiometabolic risk factors. Moreover, these previous studies focused exclusively on targeted populations, specific interventions, or single-component outcomes. Notably, Peñalvo et al. [13] investigated the effects of multicomponent workplace wellness programs on cardiometabolic health through a comprehensive meta-analysis of more than 30 years of studies published until June 2020. Their results displayed improvements in specific dietary, anthropometric, and cardiometabolic risk indicators while no definite drivers for in-between study heterogeneity such as socio-demographic, work-related or intervention characteristics were found.

Through a systematic review and meta-analysis of studies on workplace interventions targeting cardiometabolic risk factors published during the last five years, we aim to update the extant body of knowledge on this rapidly evolving field and extend previous insights by analyzing potential sources of heterogeneity and evaluating study quality with standardized assessment methods. Moreover, the results are expected to provide evidence-based recommendations for the development of future workplace health promotion programs, and to guide the

integration of cardiometabolic health promotion into broader TWH frameworks.

2. Methods

This systematic review and meta-analysis was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)-statement [14] and the Conducting Systematic Reviews and Meta-Analysis for Observational Studies of Etiology (COSMOS-E) [15] guide. The review protocol was registered on the international prospective register of systematic reviews (PROSPERO; Registration NO CRD42024617116).

2.1 Data Sources and Search Strategies

A systematic literature search was conducted in MEDLINE (PubMed), Web of Science and Embase-Ovid to identify studies published between January 1, 2019, and September 9, 2024, that evaluated the effectiveness of workplace health promotion programs targeting cardiometabolic health. A preliminary string was developed in August 2024 and initially launched on PubMed, before being adapted for the other databases. The search strings included terms such as "Health Promotion", "Health Education", "Health Campaign", "Well-being Program", and "Health Incentive Program." These were combined with workplace-related terms like "Work", "Workplace", "Occupations", "Occupational Groups", "Worker," and other terms related to cardiovascular risk factors such as "Blood Pressure", "Waist Circumference", "Body Mass Index", "Smoking Cessation", "Cholesterol", "Body Weight", "Triglycerides", "Waist-Hip Ratio", and "Blood Glucose."The complete search strings can be found in Supplementary Table 1. A research librarian was involved in the database searches to ensure methodological rigor, completeness, and accuracy.

2.2 Eligibility Criteria

Two reviewers (EP, MVP) independently screened the list of titles, abstracts and full text articles, using the Rayyan intelligent tool for systematic reviews [16]. Studies selected for full-text review were independently assessed for inclusion, with any discrepancies resolved through consensus.

The inclusion criteria were the following:

- Population: adult population at the workplace
- Intervention: single or multicomponent health promotion intervention at the workplace that targets the reduction of cardiovascular risk factors.
- Design: interventional controlled trials, including randomized controlled trials (RCTs) or quasi-experimental studies (QE).
- Publication: articles published in the last five years (between January 1, 2019, to September 9, 2024).
- Outcome: objective parameters (such as anthropometric, hematological measures, and smoking cessation) related to cardiovascular risk factors
- Effect measure: estimates of the difference in the specified outcome and a measure of uncertainty (e.g. confidence interval or standard error), or sufficient data to compute them.
- Language: studies written in English or Italian.
- A detailed summary based on the PICOS framework is presented in Supplementary Table 2.

2.3 Data Extraction

Two reviewers (EP and MVP) independently extracted relevant data from the selected papers. Extracted data was organized into five main categories: publication details, workplace characteristics, workers details, intervention characteristics and outcome measures.

Publication details included: author, publication year, geographical region, study design, use of randomization and its type (cluster or individual).

Among the workplace characteristics we extracted: work sector, number of sites involved, company size (small: <50 employees, medium: 50-249 employees, large: ≥250 employees).

Workers' details included type of control sample, job title, ISCO-08 code from the International

Labour Organization (ILO), classification as white collar and/or blue collar, mean age, predominant ethnicity, mean work seniority, education level, number of smokers and alcohol consumers, physical activity (number of sedentary and active individuals, following WHO 2020 guidelines [17] with sedentary being less than 150 minutes of moderate to intense physical activity per week), type of contract, type of work shift (day and/or night), monthly salary, health status of participants (healthy and/or affected by specific diseases/cardiovascular risk factors).

Intervention characteristics included: area of interest (single or multiple, between dietary habits, physical activity, smoking cessation, stress management, sleep hygiene, health screening, alcohol consumption reduction), and type of intervention (1. Individual communication: mobile-based/smartphone app, online lesson, interactive website, newsletter, nutritional program to follow, coach support, booklet/paper, phone call, postal letter, recurrent computer messages, sleep hygiene program, scheduled health check-ups, nicotine replacement treatment, text messages; 2. Group communication: in-person lessons, social media communication, gamification; 3. Physical activity: physical exercises; 4. Self-awareness: relaxation techniques, workplace quit smoking program, quit smoking program, stress management techniques, quit drinking program), duration of the intervention, number of interventions (total number and monthly), modality of intervention (in-person and/or online), professional figure involved (physicians: if at least one physician was involved; other healthcare professional: if at least one among nurses, nutritionists, physiotherapists, psychologists was involved; other: if the intervention was conducted by non-medical staff (e.g., sports instructors, teachers, social services, colleagues, cooking experts)), involvement of the management in the planning phase, financial incentives, and re-engagement.

Finally, the following *outcome measures*, extracted as continuous effect sizes (ES), were selected: Body Mass Index (BMI) (kg/m^2), body weight (kg), total cholesterol (mg/dL), HDL cholesterol (mg/dL), LDL cholesterol (mg/dL), triglycerides (mg/dL), systolic blood pressure (SBP) (mmHg), diastolic blood pressure (DBP) (mmHg), body fat percentage, waist circumference (cm), fasting blood glucose (FBG) (mmol/L), smoking cessation.

If possible, missing data was resolved by assumptions agreed upon by two investigators (AG and II). The full list of assumptions is available in Supplementary Table 3.

2.4 Quality of Study Assessment

The quality of the studies was independently assessed by two reviewers (EP and MVP) using a previously established scoring system [13, 18–20], which has been applied for similar works. It is based on five criteria: study design, assessment of exposure, assessment of outcome, control for confounding, and evidence of selection bias. A binary score can be attributed to each criterion (0-1). The overall score results from the sum of individual scores with 0–3 scores considered as low-quality and 4–5 considered high-quality. The detailed list of the bias assessment criteria is available in Supplementary Table 4.

2.5 Quality of Evidence Assessment

The GRADE (Grading of Recommendations Assessment, Development and Evaluation) framework was employed to evaluate the overall certainty of evidence across studies for each outcome [21]. This framework classifies the quality of evidence in systematic reviews into four levels: "high," "moderate," "low," and "very low." The initial certainty level was set high, given that most of the studies included were RCTs. The certainty of evidence was subsequently assessed for potential downgrading. Decisions regarding upgrading or downgrading are based on the criteria and considerations outlined in the GRADE handbook [22]. Reasons for downgrading include: studies' limitations, indirectness of evidence, inconsistencies across findings, imprecision, and potential publication bias. The evaluations of these criteria were primarily informed by results from the study quality assessment, measures of heterogeneity, points estimates and confidence intervals and publication bias. Criteria for upgrading are large magnitude of effect size, presence of a dose-response gradient and plausible confounding factors that reduce the effect size. The assessment

of evidence quality was conducted independently by two reviewers (MVP and EP), and discrepancies were resolved through discussion with a third reviewer (AG) to achieve consensus.

2.6 Statistical Analysis

Multiple inverse-variance random effects multilevel meta-analyses were conducted to account for dependencies between study-specific effect sizes. The multilevel approach allows for the consideration of both the variance in effect sizes within the same study (level 2) and the variance between different studies (level 3). The restricted maximum likelihood (REML) estimator was used to calculate between study heterogeneity $\tau 2$. Statistical heterogeneity was assessed using Cochran's Q test and I^2 statistics. To determine whether the more complex three-level models provided a significantly better fit to the data compared to simpler two-level models, we employed likelihood ratio (LR) tests and compared Akaike Information Criterion (AIC) values across models.

Subgroup meta-analyses were performed for outcomes with more than 10 effect sizes to explore potential sources of heterogeneity related to study design, geographic location, workplace setting, enterprise size, type of worker, presence of pre-existing health risks, intervention modality, main provider of the intervention, economic incentives, involvement of management in intervention planning and study quality. We used Knapp-Hartung adjustment [23] to reduce the risk of false significant effects. We also conducted a series of meta-regressions to investigate the association of included outcomes with participant mean age, study size, number of interventions per month, prevalence of male participants, and overall intervention duration. Multiple meta-regressions were not conducted due to an insufficient number of studies to provide reliable estimates and ensure adequate statistical power. Potential publication bias and small study effect were assessed through an adaptation of Egger's Test for multilevel meta-analysis using the standard errors as moderators and the visual inspections of funnel plots. Finally, a leave-one-out sensitivity analysis was conducted to assess the robustness of the findings by iteratively removing one study at a time and examining the impact on the overall effect estimates. All analyses were performed using RStudio (version 4.4.3) [24].

Where necessary, ESs were standardized by converting measurement units. The difference between intervention and control group changes at follow-up was either directly extracted or computed from the available data. Standard errors (SE) of the ESs were extracted or computed from available estimates whenever possible. If no relevant statistics were available, SEs were computed on the following assumptions: for paired observations without reported covariance (within-group changes at follow-up) we applied a correlation coefficient of 0.9 when loss to follow-up was below 10%, and 0.5 when loss exceeded 10%. For independent samples (between-group change at follow-up) we used a correlation coefficient of 0.

3. Results

The literature search returned 6069 articles. After removing duplicates (n = 641), 5,428 articles remained. After screening the titles and abstracts, 110 were found to be relevant for retention. Then, the full texts were examined and assessed against exclusion and inclusion criteria. Sixty-six articles did not meet the eligibility criteria. Finally, a total of 44 publications (25–68) were included, comprising 30 RCTs and 14 QE studies. Details of the search process and selection of studies are provided in Figure 1.

The full list of included articles together with their main characteristics is reported in Table 1. Most studies were conducted in Europe (n=14), and Asia (n=14), followed by North America (n=8), Middle East (n=4), South America (n = 3) and New Zealand (n=1). The occupational sectors in which the effectiveness of health promotion interventions was assessed included the tertiary sector (n=16), healthcare (n=9), industry (n=9), mixed sectors (n=6), and unspecified sectors (n=4). Concerning the number of worksites adhering to the interventions, 33 studies reported engaging between 1 and 60 sites, while 11 articles did not provide this information. The reported dimension of the enterprise involved was large for 24 studies and medium for 6 studies. Fourteen articles did not specify the company

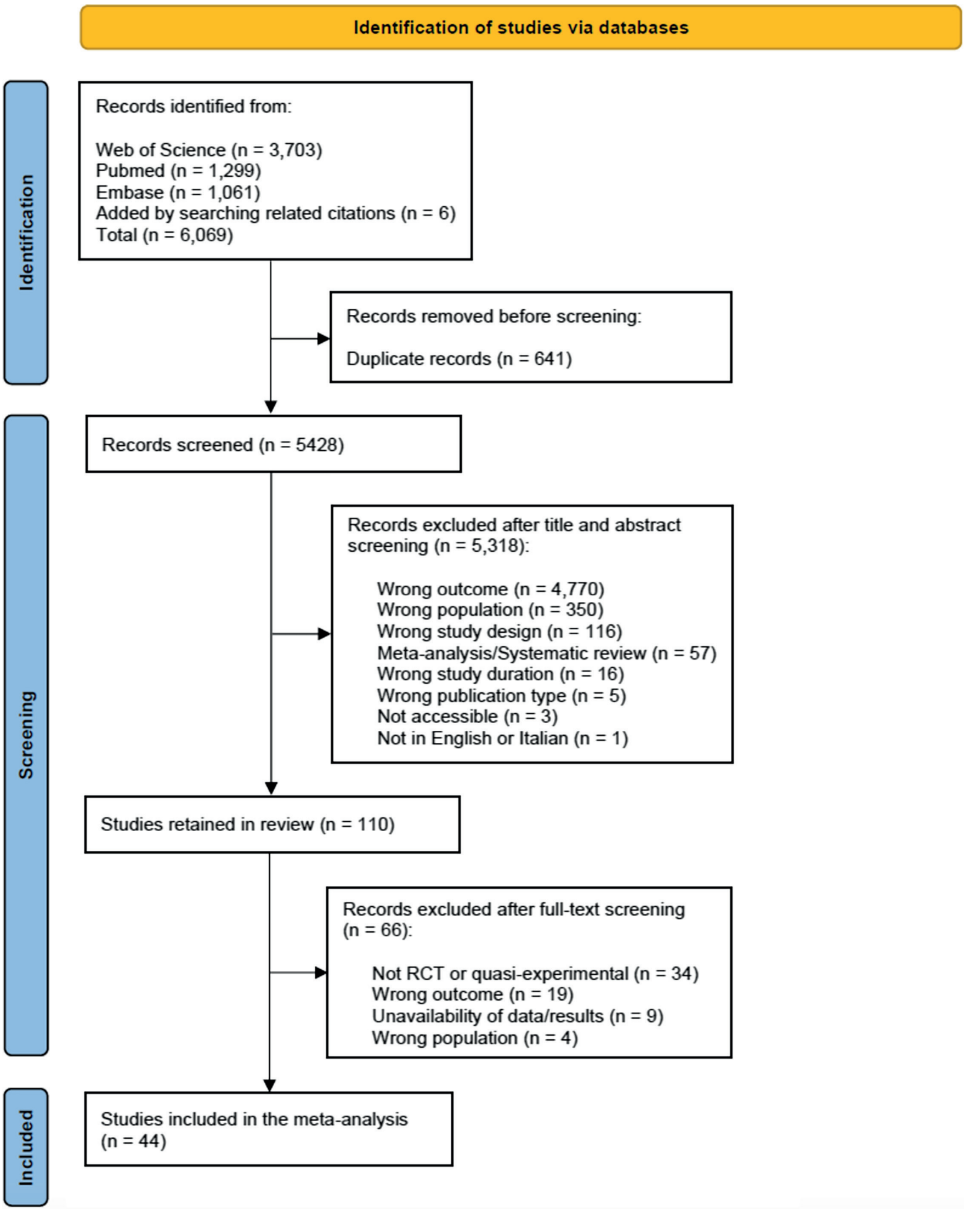

Figure 1. Study selection according to the PRISMA-flow diagram.

size. Numbers of participating employees varied considerably between studies (median 110, range 16-24396) with an overall number of participants of 49,813 (cases: 32,457, controls: 17356). The average duration of the intervention was 9.28 months (range: 1-60). Most represented areas of intervention targeted physical activity (86%, 38/44), dietary habits (48%, 21/44), followed by smoking cessation (18%, 8/44), stress management (14%, 6/44), alcohol drinking behavior (5%, 2/44) sleep hygiene (2%, 1/44), with 19 studies (43%) having more than one target. Concerning the modality of intervention, 21 (48%) adopted both web-based and in-person interventions. The most represented outcomes were BMI (n=30), body weight (n=19) and systolic blood pressure (n=20). (Supplementary Table 5). The mean quality score assessment was 4.0 (± SD 1.1) with 32 high-quality studies and 12 low-quality studies

Table 1. Main characteristics of the included studies.

First Author and year	Country	Design	Working sector	Number of employees	Number of sites	Modality of intervention	Duration (months)	Outcome	Quality Score
Arrogi et al, 2019(25)	Belgium	QE	Industry	300	4	Individual communication, Group communication, Physical exercise	9	BMI, body fat %, WC	3
Asfar et al, 2021(26)	USA	RCT	Industry	134 males	17	Individual communication, Group communication, Self-awareness	6	Smoking cessation	5
Barranco-Ruiz et al, 2019(27)	Ecuador	RCT	Tertiary	98	1	Physical exercise	4	BMI, body weight, SBP, DBP, body fat %,	4
Clemes et al, 2022(28)	UK	RCT	Tertiary	382	Unknown	Individual communication, Group communication, Physical exercise	18	BMI, weight, TOT-Chol, HDL, LDL, TG, body fat %, WC	5
Day et al, 2019(29)	USA	RCT	Tertiary	421	10	Individual communication, Group communication, Physical exercise	6	Weight	4
Diaz-Benito et al, 2021(30)	Spain	RCT	Unknown	72	Unknown	Individual communication, Group communication, Physical exercise	3	BMI, weight, WC	4
Fang et al, 2019(31)	Taiwan	RCT	Industry	75	Unknown	Group communication, Physical exercise	3	BMI, weight, TOT-Chol, HDL, LDL, TG, SBP, DBP, body fat %, WC	5
Garcia-Rojas et al, 2021(32)	Mexico	QE	Healthcare	2002	7	Individual communication, Group communication, Physical exercise	12	SBP, DBP	3
Gerodimos et al, 2021(33)	Greece	RCT	Healthcare	54	1	Group communication, Physical exercise	1,5	SBP, DBP, body fat %	5

(*continued*)

First Author and year	Country	Design	Working sector	Number of employees	Number of sites	Modality of intervention	Duration (months)	Outcome	Quality Score
Gimenez et al, 2024(34)	Brazil	RCT	Healthcare	28 females	1	Individual communication, Group communication, Physical exercise	2	BMI, weight, TOT-Chol, HDL, LDL, TG, WC, FBG	4
Guirado et al, 2024(35)	France	RCT	Tertiary	40 females	6	Group communication, Physical exercise	3	BMI, weight, TOT-Chol, HDL, LDL, TG, WC, FBG	4
Hassani et al, 2020(36)	Iran	RCT	Industry	92 males	1	Group communication, Individual communication	3	BMI, weight, body fat %, FBG	5
Hee Woo et al, 2019(37)	South Korea	RCT	Healthcare	68	Unknown	Individual communication, Group communication, Physical exercise	3	BMI, TOT-Chol, HDL, LDL, TG, SBP, DBP, WC, FBG	4
Hu et al, 2023(38)	China	RCT	Various	24396	60	Individual communication, Group communication, Physical exercise, Self-awareness, Environmental	24	SBP, DBP, smoking cessation	5
Iturriaga et al, 2019(39)	Spain	RCT	Unknown	63 females	Unknown	Group communication, Physical exercise	3	BMI, body fat %	4
Jorvand et al, 2020(40)	Iran	QE	Healthcare	114	8	Individual communication, Group communication, Physical exercise	6	TOT-Chol, HDL, LDL, TG, FBG	3
Karatrantou et al, 2020(41)	Greece	RCT	Tertiary	40	8	Group communication, Physical exercise	6	SBP, DBP, body fat %	4
Kim et al, 2023(42)	South Korea	QE	Unknown	296 males	Unknown	Group communication, Individual communication, Self-awareness	6	Smoking cessation	4
Kim et al, 2022(43)	South Korea	RCT	Various	50 females	Unknown	Individual communication, Physical exercise	6	TOT-Chol, HDL, LDL, TG, FBG	4

Koch et al, 2022(44)	Germany	RCT	Tertiary	120	1	Group communication, Physical exercise	6	SBP, DBP	5
Kong et al, 2022(45)	China	RCT	Unknown	388	2	Individual communication, Group communication, Physical exercise, Environmental	12	BMI, WC	5
Kotejoshyer et al, 2021(46)	USA	RCT	Tertiary	269	1	Individual communication, Group communication, Physical exercise	60	BMI, body fat %	5
Kugathasan et al, 2023(47)	Canada	QE	Tertiary	524	8	Individual communication, Group communication, Physical exercise	16	BMI, weight	3
Lennefer et al, 2020(48)	Germany	RCT	Industry	121	1	Individual communication, Group communication, Physical exercise	1	BMI	4
Ma et al, 2021(49)	Japan	RCT	Industry	75	1	Individual communication, Physical exercise, Environmental	3	BMI, body fat %	4
Mahdavi-Roshan et al, 2020(50)	Iran	QE	Healthcare	97 females	5	Individual communication, Group communication	2	BMI, weight, WC	4
Maphong et al, 2021(51)	Thailand	QE	Tertiary	78	2	Individual communication, Group communication, Physical exercise, Environmental	2	SBP, DBP, WC	2
Mat Azmi et al, 2022(52)	UK	QE	Tertiary	16	2	Group communication, Physical exercise, Environmental, Individual communication	2	TOT-Chol, HDL, LDL, TG, FBG	2
Moon et al, 2024(53)	South Korea	QE	Tertiary	68	2	Individual communication, Group communication, Physical exercise	3	BMI, TOT-Chol, HDL, SBP, DBP, WC, FBG	2

(*continued*)

First Author and year	Country	Design	Working sector	Number of employees	Number of sites	Modality of intervention	Duration (months)	Outcome	Quality Score
Nagata et al, 2022(54)	Japan	QE	Tertiary	3697	Unknown	Individual communication, Group communication, Environmental, Physical exercise	2	BMI, weight, LDL, SBP, DBP, WC	2
Nahm et al, 2020(55)	South Korea	RCT	Industry	40	1	Individual communication, Group communication, Physical exercise	3	BMI, weight, SBP, DBP, body fat %	4
Ozaki et al, 2019(56)	Japan	RCT	Various	80	Unknown	Individual communication, Group communication, Physical exercise	3	BMI, weight	5
Raymond et al, 2019(67)	USA	QE	Various	831	5	Individual communication, Group communication	60	BMI, TOT-Chol, HDL, LDL, TG, WC	2
Rigotti et al, 2020(68)	USA	RCT	Healthcare	106	7	Individual communication, Self-awareness	12	Smoking cessation	5
Röhling et al, 2020(57)	Germany	RCT	Healthcare	30	1	Individual communication, Group communication, Physical exercise	3	BMI, weight, TOT-Chol, HDL, LDL, TG, SBP, DBP, body fat %, WC, FBG	5
Ruettger et al, 2022(58)	UK	RCT	Tertiary	244	25	Individual communication, Group communication, Physical exercise	6	BMI, weight, TOT-Chol, HDL, LDL, TG, SBP, DBP, body fat %, WC	5

Ryu et al, 2021(59)	South Korea	QE	Tertiary	52	2	Individual communication, Physical exercise	3	BMI, TOT-Chol, HDL, LDL, TG, SBP, DBP, body fat %, WC	4
Saavedra et al, 2020(60)	Iceland	QE	Tertiary	47	1	Group communication, Physical exercise	3	BMI, weight, TOT-Chol, HDL, LDL, TG, SBP, body fat %	2
Shakerian et al, 2023(61)	Iran	RCT	Industry	588	1	Individual communication, Group communication, Physical exercise	6	BMI, body fat %	4
Song et al, 2019(62)	USA	RCT	Industry	8143	40	Individual communication, Group communication, Self-awareness	18	BMI, TOT-Chol, HDL, SBP, DBP, FBG	5
Thorndike et al, 2021(63)	USA	RCT	Healthcare	602	1	Individual communication, Environmental	24	BMI, weight, TOT-Chol, HDL, LDL, TG, SBP, DBP, WC	5
Van de Ven et al, 2023(64)	Netherlands	RCT	Various	176	Unknown	Individual communication, Group communication, Physical exercise	6	BMI, weight	5
Wang et al, 2020(65)	China	RCT	Various	4548	60	Individual communication, Environmental, Self-awareness	24	SBP, DBP, smoking cessation	5
Wilson et al, 2022(66)	New Zealand	QE	Tertiary	148	Unknown	Individual communication, Group communication, Physical exercise	4	BMI, weight, SBP, DBP, body fat %, WC	2

(Supplementary Table 6). Aggregate study characteristics are listed in Supplementary Table 7.

Pooled estimates were derived for twelve outcomes, including eight cardiovascular risk factors and four anthropometric measurements. Among the cardiovascular risk factors, blood pressure was the most frequently analyzed outcome (24 estimates from 20 studies), whereas BMI was the most examined among anthropometric measurements (39 estimates from 30 studies).

Three-level meta-analysis pooled results revealed statistically significant improvements in nine out of twelve outcomes (Table 2). All anthropometric measures showed statistically significant reductions: BMI (-0.61 kg/m2, [–0.93; –0.29]; I^2level2=63.7%, I^2level3=31.8%, p.het<0.01; Figure 2), weight (-2.42 kg, [–3.48; –1.38]; I^2level2=20.7%, I^2level3=77.2%, p.het<0.01), body fat (-1.58%, [–2.37.; –0.79]; I^2level2=15.6%, I^2level3=69.9%, p.het<0.05), and waist circumference (-3.46 cm, [–5.15; –1.76]; I^2level2=59.9%, I^2level3=36.2%, p.het<0.05). Among cardiovascular risk factors, significant changes after health promotion programs were observed for LDL cholesterol (-5.9 mg/dL, [-11.54; -0.22]; I^2level2=71.7%, I^2level3=24.7%, p.het<0.05), HDL cholesterol [2.76mg/dL, (0.41; 5.10]; I^2level2=0%, I^2level3=96.3%, p.het<0.05), DBP (-2.34 mmHg, [-3.58; -1.13]; I^2level2=1.6%, I^2level3=94.2%, p.het<0.001) and SBP (-3.746 mmHg, [-5.67; -1.83]; I^2level2=9.6%, I^2level3=83.7%, p.het<0.001; Figure 3). Finally, smoking was significantly reduced (OR: 0.79, [0.63; 0.98], I^2 =77%, p.het=0.016). No significant changes were observed for total cholesterol (-5.96 mg/dL, [-12.08; -0.92]; I^2level2=76.4%, I^2level3=18.2%, p.het<0.001); FBG (-0.98 mg/dL, [-6.44; 4.50]; I^2level2=0%, I^2level3=97.0%, p.het<0.001), triglycerides (-11.78 mg/dL, [-28.34; 4.77]; I^2level2=0%, I^2level3=96.8%, p.het<0.001) and smoking cessation (OR: 1.43, [0.99; 2.07], I^2 =88%, p.het<0.001:). The complete representation of forest plots is available in Supplementary Figures 1-10.

High within-study heterogeneity was found for BMI, weight, waist circumference, total cholesterol, and LDL cholesterol. High between-study heterogeneity was observed in the remaining outcomes, including body fat, HDL cholesterol, DBP, SBP, FBG, triglycerides, and smoking cessation. Overall high levels of heterogeneity (I^2>60%) were observed across all outcomes.

Subgroup meta-analyses and univariate meta-regressions identified significant heterogeneity (p<0.05) across several variables. Among the anthropometric outcomes, the pooled BMI estimate showed significant heterogeneity in relation to the provider of intervention (p.het=0.049), with healthcare professionals (-1.60 kg/m2 [-2.55;-0.65]) and physicians (-0.74kg/m2 [-1.25;-0.53]) achieving a more significant BMI reduction compared to other professionals (-0.39kg/m2 [-0.76;-0.02]); intervention duration (p.het=0.047), with interventions lasting less than three months proving a more effective reduction than longer ones (-0.93kg/m2 [-1.33;-0.52]), and health status of participants (p.het=0.049), with studies considering individuals with cardiovascular risk factors (-1.45 kg/m2 [-2.00; -0.90]) showing greater effectiveness compared to studies with healthy/mixed individuals (-0.38 kg/m2 [-0.67; -0.08]). The pooled estimate for weight reduction showed significant heterogeneity based on the presence of economic incentives (p = 0.002), with a greater reduction observed in studies that did not provide economic incentives (–2.99 kg [–3.99; –2.00]). Body fat reduction exhibited significant heterogeneity by geographic region (p.het=0.001), with studies conducted in Asia (–1.47% [–1.48; –0.42[), Europe (–1.27% [–0.77; –0.04]), and other countries (–1.47% [–1.48; –0.42]) showing significantly greater reductions compared to those from North America (1.33% [–0.32; 2.99]). Significant heterogeneity was also observed for intervention duration (p.het=0.003), with shorter interventions proving more effective (-2.14% [-3.08; -1.00]). For waist circumference (WC), the only significant source of heterogeneity was the provider of the intervention (p.het=0.001). Interventions led by healthcare professionals, specifically physicians (–3.58cm [–4.85; –2.32] and other health workers (–10.85cm [–14.19; –7.50]), achieved significantly greater reductions compared to those delivered by other professionals (–2.04cm [–3.72; –0.36]). Among the cardiovascular risk factors, the pooled estimate for total cholesterol showed significant heterogeneity according to the health status of participants

Table 2. Results of the multilevel random effects meta-analysis for all the included outcomes.

Outcome*			Overall estimate (95% CI)	p. Het	I^2 level 3	I^2 level 2	GRADE	
BMI (kg/m^2)	39	30	-0.61 (-0.93; -0.29)	<0.001	31.8%	67.3%	Moderate ⊕⊕⊕◯	Indirectedness
Weight (kg)	26	19	-2.42 (-3.48; -1.38)	<0.001	20.8%	77.2%	Moderate ⊕⊕⊕◯	Indirectedness
Body Fat (%)	20	16	-1.58 (-2.37; -0.79)	<0.001	62.9%	15.6%	High ⊕⊕⊕⊕	
Waist Circumference (cm)	20	17	-3.46 (-5.15; -1.76)	<0.01	36.2%	59.9%	Moderate ⊕⊕⊕◯	Indirectedness
Total Cholesterol (mg/dL)	19	16	-5.96 (-12.08; 0.92)	<0.001	18.2%	76.4%	Very low ⊕◯◯◯	Indirectedness Inconsistency Imprecision
LDL Cholesterol (mg/dL)	18	15	-5.88 (-11.54; -0.22)	<0.001	24.7%	71.7%	Very low ⊕◯◯◯	Indirectedness Inconsistency Imprecision
HDL Cholesterol (mg/dL)	19	16	2.75 (0.41; 5.10)	<0.001	96.3%	0%	Moderate ⊕⊕⊕◯	Indirectedness
DBP (mm/Hg)	22	19	2.34 (-3.58; -1.13)	<0.001	94.2%	1.6%	High ⊕⊕⊕⊕	
SBP (mm/Hg)	24	20	-3.75 (-5.67; -1.83)	<0.001	83.7%	9.6%	High ⊕⊕⊕⊕	
FBG (mmol/L)	12	11	-0.98 (-6.44; 4.50)	<0.001	97.0%	0%	Low ⊕⊕◯◯	Inconsistency Imprecision
Triglycerides (mg/dL)	17	11	-11.79 (-27.34; 3.75)	<0.001	96.8%	0%	Very low ⊕◯◯◯	Indirectedness Inconsistency Imprecision
Smoking (OR)	5	5	0.79 (0.63; 0.98)	0.016	77.0 %		Low ⊕⊕◯◯	Indirectedness Publication bias

Anthropometric outcomes include BMI, weight, body fat, and waist circumference (WC). Cardiovascular risk factors include total cholesterol (TOT-Chol), systolic blood pressure (SBP), diastolic blood pressure (DBP), fasting blood glucose (FBG), and smoking reduction.GRADE level for certainty of evidence: 'high' indicates that we are very confident that the true effect lies close to that of the estimate of the effect; 'moderate' indicates that the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different; 'low' indicates that the true effect may be substantially different from the estimate of the effect; and 'very low' indicates that the true effect is likely to be substantially different.*

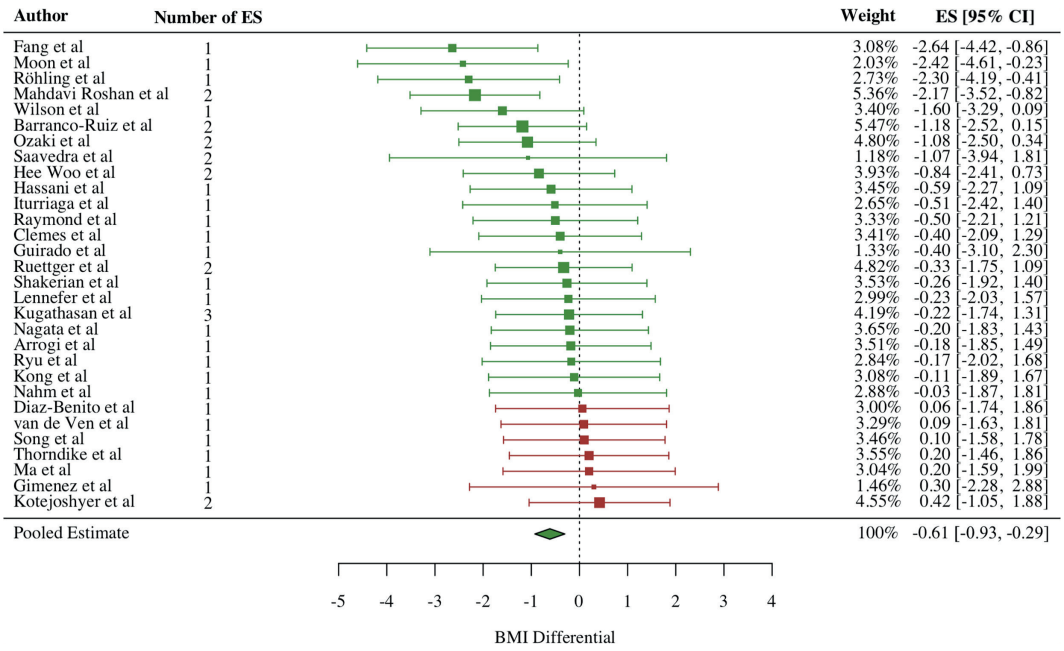

Figure 2. Forest plot of BMI (kg/m2).

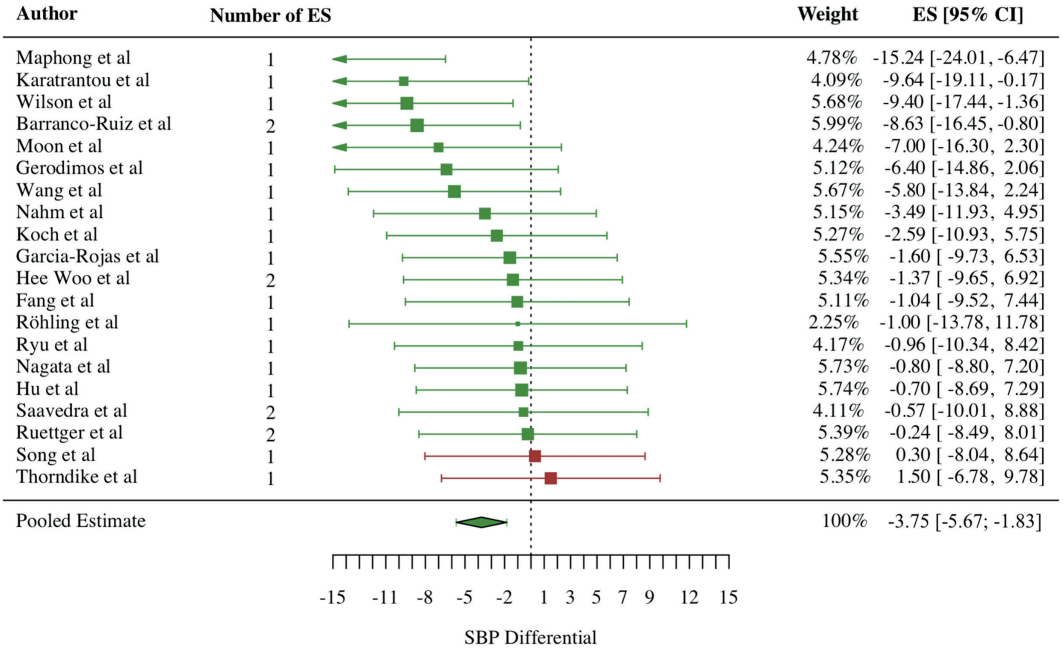

Figure 3. Forest plot of SBP (mm/Hg).

(p.het=0.014), with studies involving individuals with cardiovascular risk factors showing a greater reduction (–22.52mg/dL [–36.59; –8.63]) compared to those involving healthy individuals (–2.52mg/dL [–8.93; 3.80]). The same result was obtained for LDL cholesterol (p.het=0.001) with a reduction of (–24.17mg/dL [–33.09; –15.24]) for individuals with cardiovascular risk factors. The pooled results

for HDL cholesterol showed significant heterogeneity based on enterprise size (p.het=0.001), with studies involving medium-sized enterprises reporting a significant increase in HDL cholesterol (13.83mg/dL [9.01; 18.66]). No significant sources of heterogeneity were identified for FBG and DBP. For SBP, heterogeneity was observed only in relation to the presence of economic incentives, with a significant reduction found in studies that did not offer such incentives (–4.76mmHg [–6.77; –2.76]). Finally, triglycerides pooled estimates showed significant heterogeneity according to enterprise size (p.het=0.015), with medium-sized enterprises again demonstrating a significant decrease (–42.29mg/dL [–70.98; –13.62]). Among the meta-regressions, we found a significant association between the number of interventions per month and LDL cholesterol (β=0.12, p=0.025). Another association was found between the mean age of participants and FBG (β = 0.46, p = 0.000). Full results of the subgroup meta-analyses and meta-regressions are available in Supplementary Table 8.

A visual inspection of funnel plots and an adaptation of Egger's regression tests was performed to evaluate potential small-study effects or publication bias (Supplementary Figures 11-22). Egger's test indicated potential small-study effects or publication bias for triglycerides (p=0.037), and smoking cessation (p=0.006), which was further supported by the asymmetry observed in the corresponding funnel plot. Visual inspection of the funnel plots suggested a potential asymmetry for LDL cholesterol, and FBG although Egger's test was not statistically significant. No asymmetry in the plots or significant results from Egger's tests were observed for several outcomes, including BMI, weight, body fat, total cholesterol, HDL, LDL, waist circumference, DBP, and SBP.

In the leave-one-out sensitivity analysis, similar pooled effects were observed across all anthropometric outcomes, indicating the robustness of these results. However, for cardiovascular risk factors, the exclusion of specific studies affected the statistical significance of some pooled estimates. The total cholesterol estimate became significant upon the removal of Ryu et al. [59], Raymond et al. [67], and Gimenez et al. [34]; FBG estimate became significant after excluding the study by Fang et al. [31]; and triglycerides reached significance after excluding the study by Kim et al. [42]. Conversely, the pooled estimate for LDL cholesterol lost its statistical significance when the study by Fang et al. [31] was removed.

The quality of evidence, as assessed using the GRADE system, is presented in Table 2, with a detailed justification for each rating available in Supplementary Table 9.

4. Discussion

This systematic review and meta-analysis synthesized the evidence from 44 studies assessing the effectiveness of workplace-based health promotion interventions on cardiometabolic health outcomes. The main finding of our analysis is that such interventions can lead to significant improvements across a wide range of anthropometric and cardiovascular risk parameters. Specifically, we observed significant improvements in nine out of twelve cardiometabolic outcomes. Although changes in individual parameters were generally modest, even small improvements can lead to meaningful health benefits at both the individual and population levels. Evidence shows that slight, but sustained changes in these parameters directly contribute to reducing cardiovascular and cerebrovascular events, decreasing the risk of chronic metabolic diseases, and reducing systemic inflammation [69,70]. Moreover, cardiometabolic risk is deeply intertwined with mental health, psychological well-being, and overall quality of life [71–73]. In addition to direct health benefits, these improvements can yield indirect advantages for employers and healthcare systems. Weight loss, improved blood pressure control, and better lipid profiles are associated with reduced absenteeism, greater productivity, and job satisfaction potentially contributing to preserved work capacity and extended working life [74–76]. Improved cardiometabolic health is linked to lower disease burden, treatment costs, and decreased resource utilization, leading to important implications for public health and healthcare systems [77].

These results support the growing body of evidence that the workplace is an effective setting for

the implementation of multidimensional health promotion strategies. In this regard, a meta-analysis by Peñalvo et al. [20] evaluated the impact of multicomponent workplace interventions on dietary habits, overweight, and cardiometabolic health, by analyzing 121 studies conducted between 1990 and 2020. The authors highlighted a significant increase in fruit and vegetable consumption and HDL cholesterol, and significant reductions in BMI, body weight, SBP, DBP, LDL cholesterol, triglycerides, and FBG. Other previous meta-analyses [73, 78–80] with a lower number of included studies and mainly focused on lifestyle interventions and dietary habits yielded similar promising results.

Despite these positive findings, the observed between-study heterogeneity was consistently high across all outcomes, possibly due to variations in study design, sociodemographic characteristics, workers' details, and implementation contexts. Subgroup analyses and meta-regressions explained some of this heterogeneity. For instance, interventions led by healthcare professionals, especially physicians, were more effective in reducing BMI and waist circumference. This aligns with results presented by Zusman et al. [81], highlighting the impact of the provider of the intervention and the need to match their clinical expertise with the proposed intervention and the desired outcome. Additionally, shorter interventions [<3 months] were associated with greater improvements in BMI and body fat. Shorter lifestyle interventions tend to achieve higher adherence, as maintaining motivation and consistent behavioral change is easier over limited periods. This may also relate to a novelty effect, whereby enthusiasm and commitment are strongest early on. Another possible explanation is selection bias, as participants in longer programs could be more prone to drop out if early results are not achieved. Similar results were found in a recent systematic review and meta-analysis by Rotunda et al. [82] investigating the effectiveness of lifestyle interventions lasting 6 months or less on the body weight of adults with overweight or obesity, concluding that interventions lasting less than 13 weeks were at least as effective as longer ones [13-26 weeks]. Early phases of intervention often yield the greatest weight loss, and shorter multicomponent programs tend to have higher adherence and compliance resulting in a greater retention rate [83]. Moreover, early weight loss has been identified as a predictor of greater long-term weight reduction [84]. Baseline cardiometabolic profile also emerged as a possible moderator, with employees already at higher cardiovascular risk benefiting more from interventions targeting BMI, total cholesterol, and LDL cholesterol. This finding aligns with previous studies [85, 86] highlighting better results in high-risk populations. With regard to unexplained heterogeneity, inconsistent reporting of certain variables across studies limited our ability to explore key sources of variation. Most studies did not report baseline cardiovascular-related characteristics, such as dietary habits or physical activity, nor participants' socioeconomic status. Information on work schedules, including shift or night work, was also generally missing, along with other occupational risk factors such as workload and job stress. Few studies provided details on the engagement of workers in program planning, despite its potential impact on participation and motivation. Further sources of heterogeneity are likely contextual, with multiple layers potentially influencing the effectiveness of occupational health promotion programs, including factors such as country, culture, language, corporate culture, job roles, and organizational implementation. The awareness of the sources of heterogeneity is essential to drive future health promotion programs. Factors influencing effectiveness – such as the workforce's cardiometabolic profile, the intervention provider and the duration of the programs – should be carefully considered to optimize cardiometabolic outcomes. In this regard, occupational physicians play a crucial role, given their expertise in both the health impacts of work environments, exposures and organization and individual susceptibility factors. This comprehensive perspective allows them to support employers and policymakers to develop integrated, tailored health strategies that align with enterprise characteristics and workers' specific health and safety needs. The increasing availability of digital health technologies (e.g., mobile apps, telehealth, wearables) may further enhance the scalability of these programs, reducing barriers related to geographic and resource constraints. It's worth noting that most of the

analyzed studies focused only on traditional workplace health promotion programs rather than adopting integrated and holistic approaches, such as the TWH model. Future preventive strategies should also tackle organizational and environmental factors to promote both healthier workplaces and healthier individual behaviors. TWH builds on the recognition that work is a social determinant of health and seeks to improve workers' health and well-being by targeting working conditions and individual factors, thereby reducing their possible additive effect [87].

We believe that the results of our meta-analysis may have important implications for public and occupational health practices and policymaking. The workplace represents a unique, yet underutilized, setting for the implementation of preventive strategies, reaching a large proportion of the adult population during their most productive year. Integrating structured health promotion interventions into occupational health policies could contribute to the reduction of the non-communicable disease burden.

Our study has several strengths. Multiple cardiometabolic outcomes were considered, along with their drivers, through a methodologically sound approach enabling a comprehensive analysis of the factors associated with cardiometabolic health outcomes. Additionally, it offers a pragmatic contribution to cardiovascular health promotion by focusing on practical aspects of the initiatives that can inform the development of future effective strategies. Furthermore, the study includes articles published after 2020, a period marked by the transformative impact of the COVID-19 pandemic, which significantly affected work patterns, efficiency, and productivity, and cardiometabolic health [90]. The long-term consequences on worker well-being and cardiometabolic profile are still unknown and unfolding, underscoring the importance of adapting health promotion interventions to the new post-pandemic work environments. Finally, unlike previous published meta-analyses, we adopted a validated tool (GRADE) to assess the certainty of evidence.

Several limitations should also be acknowledged. The study revealed significant between-study heterogeneity for most of the outcomes, which was only partially explained by subgroup meta-analyses. Some variables were not consistently reported across studies, reducing our ability to examine potentially important sources of variation. As a result, conclusions should be considered carefully, recognizing that the unexplained variability may influence the magnitude of the pooled effects. Furthermore, the certainty of evidence for several outcomes, including total cholesterol, LDL cholesterol, FBG, triglycerides and smoking cessation, was rated as "low" or "very low" according to the GRADE framework, indicating high uncertainty regarding the true effect estimates. These results did not remain consistent in sensitivity analyses and should be therefore interpreted with caution, in contrast to other cardiometabolic parameters supported by moderate or high-certainty evidence. Additionally, publication bias or small study effects were detected for triglycerides and smoking cessation. Consequently, the generalizability and reliability of these findings should be cautious. Moreover, a wide range of modalities of interventions was considered, both single-component and multi-component, making it difficult to isolate the effect of specific components of the health promotion programs. The follow-up duration was generally under 12 months and in most cases without re-engagement, limiting the assessment of long-term effectiveness and possibly overestimating short-term benefits. Therefore, it is important to interpret our results as evidence of short- to medium-term effectiveness, acknowledging that the long-term sustainability of these benefits remains unclear. Long-term data are needed to determine whether initial improvements are maintained beyond the intervention period. This represents a critical knowledge gap that future research should address through extended follow-up assessments and periodic re-engagement strategies. Finally, most of the studies were conducted in Europe and Asia, potentially affecting the generalizability of our findings.

Our findings support the inclusion of workplace-based health promotion programs within national and global public health strategies, such as the EU Healthier Together Initiative [88] and the WHO Global Action Plan for the Prevention and Control of Noncommunicable Diseases [89]. To this end, policymakers should consider some essential actions, including (i) encouraging cross-sector collaboration

among stakeholders: healthcare providers, enterprises, public institutions and academia; (ii) supporting the implementation of the TWH model within occupational health and safety frameworks; (iii) assessing workers needs in terms of safety and health to define suitable preventive measures and health promotion strategies; (iv) offering fiscal or accreditation incentives to enterprises that implement evidence-based health promotion programs.

In this perspective, workplace health promotion should be recognized not only to enhance individual well-being, but also as a strategic tool to reduce health inequalities, strengthen workforce resilience, and support sustainable economic growth.

Further studies should prioritize longer follow-up durations and incorporate periodic employees' re-engagement to provide more insights into the durability of the effects. Future research should also explore the optimal frequency, intensity, and combination of intervention components to identify the most effective strategies for improving cardiometabolic health. Moreover, integrated approaches combining individual-level interventions with organizational and environmental changes in line with the TWH model are needed. Lastly, future studies should incorporate implementation science frameworks to assess barriers and facilitators influencing occupational health promotion program adoption, scalability, and long-term sustainability across diverse sectors, workplace settings, and employees' populations.

5. Conclusion

In conclusion, this systematic review and meta-analysis provides evidence that workplace-based health promotion interventions can lead to significant improvements in cardiometabolic health outcomes. Given the workplace's unique position to reach a large and diverse adult population, integrating structured health promotion into occupational health policies offers a promising strategy to improve occupational and public health, reduce healthcare costs, and support workforce productivity. Although the observed changes are generally modest, they have the potential to reduce the burden of cardiovascular and metabolic diseases at the population level. However, the short follow-up durations and partly unexplained heterogeneity across studies warrant caution in interpreting the findings and limit conclusions on long-term effectiveness. Future research should aim to optimize intervention designs, extend follow-up periods, and adopt integrated approaches in line with the TWH approach to maximize long-term benefits and sustainability.

Supplementary Materials: The following are available online: Figure S1-10: Forest plots, Figure S11-22: Funnel plots, Table S1: Search strings on different electronic databases: PubMed, Embase, Web of Science, Table S2: PICOS framework – inclusion and exclusion criteria, Table S3: Study assumptions for SEs and Ess calculation, Table S4: Quality assessment criteria, Table S5: Numbers of studies investigating different outcomes, Table S6: Quality assessment of included studies, Table S7: Aggregate characteristics of included studies, Table S8: Stratified meta-analyses and univariate meta-regressions results from three-levels random effects models, Table S9: Summary of GRADE ratings and justifications for downgrading, Table S10: PRISMA Checklist.

Funding: This research was partially funded by the European Union Horizon Europe Research and Innovation Programme under the Cancer Prevention at Work (CPW) Project Grant Agreement No. 101104716.

Acknowledgments: Authors extend appreciation to the members of the CPW:

- Ethics Advisory Board with Legal and Ethics Manager: Anna Lo Monaco, Pietro Apostoli, Jana Holcikova, Morten Sodemann
- Multi-stakeholder Advisory Board: Francesco Ciabuschi, Jennifer Takhar, Susana Viegas
- Risk Advisory Board: Eva Negri Pedro Arcos Gonzalez, Lavinia Calugareanu, Zora Kľocová Adamčáková

for their invaluable support and guidance throughout the project's execution. for their invaluable support and guidance throughout the project's execution.

Declaration of Interest: The authors declare no conflict of interest.

Author Contribution Statement: Conceptualization, A.G., V.L., and I.I.; methodology, M.V.P and P.B; validation, V.L. and C.C.; formal analysis, M.V.P. and P.B.; data

curation, A.G., E.P., E.C. and M.V.P.; writing—original draft preparation, A.G., E.C., M.V.P. and V.L.; writing—review and editing, A.S-K., P.B., F.V., F.B. and I.I.; supervision, A.S-K., P.B., F.V., F.B. and I.I.; funding acquisition, A.G.; A.S-K., C.C., and P.B. All authors have read and agreed to the published version of the manuscript.

Declaration on the use of AI: None.

References

1. World Health Organization (WHO). Constitution of The World Health Organization. 1946. Available at https://apps.who.int/gb/bd/PDF/bd47/EN/constitution-en.pdf?ua=1.
2. World Health Organization (WHO). Available Ottawa charter for health promotion. 1986. Available at: https://www.who.int/publications/i/item/ottawa-charter-for-health-promotion.
3. World Health Organization (WHO). Closing the gap in a generation: health equity through action on the social determinants of health - Final report of the Commission on Social Determinants of Health. 2008. Available at: https://www.who.int/publications/i/item/WHO-IER-CSDH-08.1.
4. International Conference on Primary Health Care. Declaration of Alma-Ata. *WHO Chron*. 1978 Nov;32(11):428–30.
5. World Bank Group. Labor force, Total. 2025. Available at https://data.worldbank.org/indicator/SL.TLF.TOTL.IN.
6. DyvikHE.Employmentrateworldwide2000-2025.2025. Available at: https://www.statista.com/study/117978/employment-worldwide/.
7. Iavicoli I, Spatari G, Chosewood LC, Schulte PA. Occupational Medicine and Total Worker Health®: from preventing health and safety risks in the workplace to promoting health for the total well-being of the worker. *Med Lav*. 2022;113(6):e2022054.
8. Tamers SL, Streit J, Pana-Cryan R, et al. Envisioning the future of work to safeguard the safety, health, and well-being of the workforce: A perspective from the CDC's National Institute for Occupational Safety and Health. *Am J Ind Med*. 2020;63(12):1065–84.
9. Sullivan PW, Ghushchyan V, Wyatt HR, Hill JO. The medical cost of cardiometabolic risk factor clusters in the United States. *Obesity* (Silver Spring). 2007;15(12):3150–8.
10. Hwang WJ, Kang SJ. Interventions to Reduce the Risk of Cardiovascular Disease among Workers: A Systematic Review and Meta-Analysis. *Int J Environ Res Public Health*. 2020;17(7):2267.
11. Mulchandani R, Chandrasekaran AM, Shivashankar R, et al. Effect of workplace physical activity interventions on the cardio-metabolic health of working adults: systematic review and meta-analysis. *Int J Behav Nutr Phys Act*. 2019;16(1):134.
12. Park SH, Kim SY. Effectiveness of worksite-based dietary interventions on employees' obesity: a systematic review and meta-analysis. *Nutr Res Pract*. 2019;13(5):399.
13. Peñalvo JL, Sagastume D, Mertens E, et al. Effectiveness of workplace wellness programmes for dietary habits, overweight, and cardiometabolic health: a systematic review and meta-analysis. *Lancet Public Health*. 2021;6(9):e648–60.
14. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;n71.
15. Dekkers OM, Vandenbroucke JP, Cevallos M, Renehan AG, Altman DG, Egger M. COSMOS-E: Guidance on conducting systematic reviews and meta-analyses of observational studies of etiology. *PLoS Med*. 2019 21;16(2):e1002742.
16. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Syst Rev*. 2016;5(1):210.
17. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med*. 2020;54(24):1451–62.
18. Shangguan S, Afshin A, Shulkin M, et al. A Meta-Analysis of Food Labeling Effects on Consumer Diet Behaviors and Industry Practices. *Am J Prev Med*. 2019;56(2):300–14.
19. Micha R, Wallace SK, Mozaffarian D. Red and processed meat consumption and risk of incident coronary heart disease, stroke, and diabetes mellitus: a systematic review and meta-analysis. *Circulation*. 2010;121(21):2271–83.
20. Afshin A, Peñalvo JL, Del Gobbo L, et al. The prospective impact of food pricing on improving dietary consumption: A systematic review and meta-analysis. *PLoS One*. 2017;12(3):e0172277.
21. Xie CX, Machado GC. Clinimetrics: Grading of Recommendations, Assessment, Development and Evaluation (GRADE). *J Physiother*. 2021;67(1):66.
22. Brozek JL, Canelo-Aybar C, Akl EA, et al. GRADE Guidelines 30: the GRADE approach to assessing the certainty of modeled evidence-An overview in the context of health decision-making. *J Clin Epidemiol*. 2021 Jan;129:138–50.
23. Röver C, Knapp G, Friede T. Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. *BMC Med Res Methodol*. 2015;15(1):99.
24. RStudio Team. RStudio: Integrated Development for R. (Internet). Boston, MA; 2020 (cited 2025 Mar 20). Available from: http://www.rstudio.com/.
25. Arrogi A, Schotte A, Bogaerts A, Boen F, Seghers J. Increasing employees' health by workplace physical activity counseling: The mediating role of step-based

physical activity behavior change. *J Phys Act Health*. 2019;16(3):205–13.
26. Asfar T, Arheart KL, McClure LA, et al. Implementing a Novel Workplace Smoking Cessation Intervention Targeting Hispanic/Latino Construction Workers: A Pilot Cluster Randomized Trial. *Health Educ Behav*. 2021;48(6):795–804.
27. Barranco-Ruiz Y, Ramírez-Vélez R, Martínez-Amat A, Villa-González E. Effect of two choreographed fitness group-workouts on the body composition, cardiovascular and metabolic health of sedentary female workers. *Int J Environ Res Public Health*. 2019;16(24).
28. Clemes SA, Varela-Mato V, Bodicoat DH, et al. The effectiveness of the Structured Health Intervention For Truckers (SHIFT): a cluster randomised controlled trial (RCT). *BMC Med*. 2022;20(1).
29. Day RS, Jahnke SA, Haddock CK, Kaipust CM, Jitnarin N, Poston WSC. Occupationally Tailored, Web-Based, Nutrition and Physical Activity Program for Firefighters Cluster Randomized Trial and Weight Outcome. *J Occup Environ Med*. 2019;61(10):841–8.
30. Díaz-Benito VJ, Barriopedro Moro MI, et al. Effects of worksite exercise intervention (PRODET®) on well-being at work and capability in performing work-related sedentary tasks: A pilot study. *Work*. 2022;72(3): 909–20.
31. Fang YY, Huang CY, Hsu MC. Effectiveness of a physical activity program on weight, physical fitness, occupational stress, job satisfaction and quality of life of overweight employees in high-tech industries: a randomized controlled study. *JOSE*. 2019;25(4):621–9.
32. Garcia-Rojas IJ, Omidakhsh N, Arah OA, Krause N. Blood Pressure Changes After a Health Promotion Program Among Mexican Workers. *Front Public Health*. 2021;9.
33. Gerodimos V, Karatrantou K, Papazeti K, Batatolis C, Krommidas C. Workplace exercise program in a hospital environment: an effective strategy for the promotion of employees physical and mental health. A randomized controlled study. *Int Arch Occup Environ Health*. 2022;95(7):1491-1500.
34. Gimenez LB, Teixeira do Amaral V, Locato G, Marçal IR, Ferron AJT, Ciolac EG. Gamification as a Tool for Promoting Physical Exercise and Healthy Eating Habits in Healthcare Worker Women: Effects on Cardiometabolic Health and Physical Fitness at Workplace. *Am J Health Promot*. 2024;38(6):820–4.
35. Guirado T, Metz L, Pereira B, , et al. A 12-Week Cycling Workstation Intervention Improves Cardiometabolic Risk Factors in Healthy Inactive Office Workers. *J Occup Environ Med*. 2022;64(8):E467–74.
36. Hassani B, Amani R, Haghighizadeh MH, Araban M. A priority oriented nutrition education program to improve nutritional and cardiometabolic status in the workplace: A randomized field trial. *J Occup Med Toxicol*. 2020;15(1).
37. Hee Woo SH, Oh EG, Kim KS, Chu SH, Kim GS, Nam CM. Development and Assessment of a Social Network Service-Based Lifestyle-Modification Program for Workers at High Risk of Developing Cardiovascular Disease. *Workplace Health Saf*. 2020;68(3):109–20.
38. Hu Z, Wang X, Hong C, , et al. Workplace-based primary prevention intervention reduces incidence of hypertension: a post hoc analysis of cluster randomized controlled study. *BMC Med*. 2023;21(1).
39. Iturriaga T, Barcelo O, Diez-Vega I, et al. Effects of a short workplace exercise program on body composition in women: A randomized controlled trial. *Health Care Women Int*. 2020;41(2):133–46.
40. Jorvand R, Ghofranipour F, Haerimehrizi A, Tavousi M. Evaluating the impact of HBM-based education on exercise among health care workers: The usage of mobile applications in Iran. *BMC Public Health*. 2020;20(1).
41. Karatrantou K, Gerodimos V, Manouras N, et al. Health-Promoting Effects of a Concurrent Workplace Training Program in Inactive Office Workers (HealPWorkers): A Randomized Controlled Study. *Am J Health Promot*. 2020;34(4):376–86.
42. Kim W, Kim AR, Ock M, et al. Effects of a supportive workplace environment on the success rate for smoking cessation camp. *Ann Occup Environ Med*. 2023;35(1).
43. Kim Y, Lee H, Chung ML. Living labs for a mobile app-based health program: effectiveness of a 24-week walking intervention for cardiovascular disease risk reduction among female Korean-Chinese migrant workers: a randomized controlled trial. *Archives of Public Health*. 2022;80(1).
44. Koch S, Esch T, Werdecker L. Effects of a Yoga-Based Stress Intervention Program on the Blood Pressure of Young Police Officers: A Randomized Controlled Trial. *J Integr Complement Med*. 2022;28(3):234–40.
45. Kong J, Chen Y, Zheng Y, et al. Effectiveness of a Worksite-Based Lifestyle Intervention on Employees' Obesity Control and Prevention in China: A Group Randomized Experimental Study. *Int J Environ Res Public Health*. 2022;19(11).
46. Kotejoshyer R, Gilmer DO, Namazi S, Farr D, Henning RA, Cherniack M. Impact of a total worker health® mentoring program in a correctional workforce. *Int J Environ Res Public Health*. 2021;18(16).
47. Kugathasan TA, Gilbert JA, Laberge S, Tremblay J, Mathieu ME. Activate Your Health: impact of a real-life programme promoting healthy lifestyle habits in Canadian workers. *Health Promot Int*. 2023;38(3).
48. Lennefer T, Lopper E, Wiedemann AU, Hess U, Hoppe A. Improving Employees' Work-Related Well-Being and Physical Health Through a Technology-Based Physical Activity Intervention: A Randomized Intervention- Control Group Study. *J Occup Health Psychol*. 2020;25(2):143-158
49. Ma J, Ma D, Li Z, Kim H. Effects of a workplace sit–stand desk intervention on health and productivity. *Int J Environ Res Public Health*. 2021;18(21).

50. Mahdavi-Roshan M, Salimi S, Pourghane P, Ashouri A, Haghighatkhah M, Karami S. A comparative study of the effectiveness of self-management and group management on the amount of weight loss of nurses under low-calorie diet treatment: A simultaneous mixed-methods study. *J Educ Health Promot*. 2024;13(1).
51. Maphong R, Nakhonket K, Sukhonthasab S. The effectiveness of two levels of active office interventions to reduce sedentary behavior in office workers: a mixed-method approach. *Arch Environ Occup Health*. 2022;77(6):504–13.
52. Mat Azmi ISM, Wallis GA, White MJ, Puig-Ribera A, Eves FF. Desk based prompts to replace workplace sitting with stair climbing; a pilot study of acceptability, effects on behaviour and disease risk factors. *BMC Public Health*. 2022;22(1).
53. Moon JH, Ryu H. Salutogenesis intervention improves cardio-cerebrovascular health in at-risk office workers: A quasi-experimental study. *Public Health Nurs*. 2024;41(4):690–703.
54. Nagata H, Sato K, Haseda M, Kobayashi Y, Kondo N. A novel behavioral science-based health checkup program and subsequent metabolic risk reductions in a workplace: Checkup championship. *Prev Med* (Baltim). 2022;164.
55. Nahm JW, Shin YJ. Effects of mobile-based exercise intervention on health indices by the comparison of personal training time in male workers. Vol. 34. 2020.
56. Ozaki I, Watai I, Nishijima M, Saito N. Randomized controlled trial of Web-based weight-loss intervention with human support for male workers under 40. *J Occup Health*. 2019;61(1):110–20.
57. Röhling M, Martin K, Ellinger S, Schreiber M, Martin S, Kempf K. Weight reduction by the low-insulin-method— a randomized controlled trial. *Nutrients*. 2020;12(10):1–17.
58. Ruettger K, Clemes SA, Chen YL, et al. Drivers with and without Obesity Respond Differently to a Multi-Component Health Intervention in Heavy Goods Vehicle Drivers. *Int J Environ Res Public Health*. 2022;19(23).
59. Ryu H, Jung J, Moon J. Effectiveness of a Mobile Health Management Program With a Challenge Strategy for Improving the Cardiovascular Health of Workers. *J Occup Environ Med*. 2021;63(3):E132–7.
60. Saavedra JM, Kristjánsdóttir H, Gunnarsson SB, García-Hermoso A. Effects of 2 physical exercise programs (circuit training and brisk walk) carried out during working hours on multidimensional components of workers' health: A pilot study. *Int J Occup Med Environ Health*. 2020;34(1):39–51.
61. Shakerian B, Mohammad RF, Saeid D, et al. Improving Physical Fitness and Health of Office Workers in Iran. *Nurs Sci Q*. 2023;36(2):186–93.
62. Song Z, Baicker K. Effect of a Workplace Wellness Program on Employee Health and Economic Outcomes: A Randomized Clinical Trial. JAMA – *J Am Med Assoc*. 2019;321(15):1491–501.
63. Thorndike AN, McCurley JL, Gelsomin ED, et al. Automated behavioral workplace intervention to preventweight gain and improve diet. *JAMA Netw Open*. 2021;4(6):E2112528–E2112528.
64. van de Ven D, Schuring M, Kouwenhoven-Pasmooij TA, et al. Reach and effectiveness of a worksite health promotion program combining a preventive medical examination with motivational interviewing; a quasi-experimental study among workers in low socioeconomic position. *BMC Public Health*. 2023;23(1).
65. Wang Z, Wang X, Shen Y, et al. Effect of a Workplace-Based Multicomponent Intervention on Hypertension Control: A Randomized Clinical Trial. *JAMA Cardiol*. 2020;5(5):567–75.
66. Wilson D, Driller M, Winwood P, Clissold T, Johnston B, Gill N. The Effectiveness of a Combined Healthy Eating, Physical Activity, and Sleep Hygiene Lifestyle Intervention on Health and Fitness of Overweight Airline Pilots: A Controlled Trial. *Nutrients*. 2022;14(9).
67. Raymond LW, Roy DM, Mullinax SL, Yanni A, Pentek KC, Isaacs SE. Preventing Diabetes in the Workplace: Effects of Coaching and Monetary Incentives. *J Occup Environ Med*. 2019;61(7):E308–11.
68. Rigotti NA, Kelley JHK, Regan S, et al. Enhancing employer coverage of smoking cessation treatment: A randomized trial of the Partners in Helping You Quit (PiHQ) program. *Prev Med* (Baltim). 2020;140.
69. Tejada B, Joehanes R, Hwang SJ, et al. Systemic Inflammation is Associated with Cardiometabolic Risk Factors and Clinical Outcomes. *J Inflamm Res*. 2022; 15:6891–903.
70. Global Cardiovascular Risk Consortium et al. Global Effect of Modifiable Risk Factors on Cardiovascular Disease and Mortality. *New England Journal of Medicine*. 2023;389(14):1273–85.
71. Magallares A, Pais-Ribeiro JL. Mental Health and Obesity: A Meta-Analysis. *Appl Res Qual Life*. 2014; 9(2):295–308.
72. Martinelli LMB, Mizutani BM, Mutti A, D'elia MPB, Coltro RS, Matsubara BB. Quality of life and its association with cardiovascular risk factors in a community health care program population. *Clinics* (Sao Paulo). 2008;63(6):783–8.
73. Sullivan PW, Ghushchyan V, Wyatt HR, Wu EQ, Hill JO. Impact of cardiometabolic risk factor clusters on health-related quality of life in the U.S. *Obesity* (Silver Spring). 2007;15(2):511–21.
74. Bilger M, Finkelstein EA, Kruger E, Tate DF, Linnan LA. The effect of weight loss on health, productivity, and medical expenditures among overweight employees. *Med Care*. 2013;51(6):471–7.
75. Unmuessig V, Fishman PA, Vrijhoef HJM, Elissen AMJ, Grossman DC. Association of Controlled and Uncontrolled Hypertension With Workplace Productivity. *J Clin Hypertens* (Greenwich). 2016;18(3):217–22.

76. Sullivan PW, Ghushchyan V, Wyatt HR, Wu EQ, Hill JO. Productivity costs associated with cardiometabolic risk factor clusters in the United States. *Value Health*. 2007;10(6):443–50.
77. Caro JJ, O'Brien JA, Hollenbeak CS, et al. Economic Burden and Risk of Cardiovascular Disease and Diabetes in Patients with Different Cardiometabolic Risk Profiles. *Value in Health*. 2007;10:S12–20.
78. Hwang WJ, Kang SJ. Interventions to Reduce the Risk of Cardiovascular Disease among Workers: A Systematic Review and Meta-Analysis. *Int J Environ Res Public Health*. 2020;17(7):2267.
79. Bezzina A, Clarke ED, Ashton L, Watson T, James CL. Workplace Health Promotion Programs Targeting Smoking, Nutrition, Physical Activity, and Obesity in Men: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Health Educ Behav*. 2024;51(1):113–27.
80. Groeneveld IF, Proper KI, van der Beek AJ, Hildebrandt VH, van Mechelen W. Lifestyle-focused interventions at the workplace to reduce the risk of cardiovascular disease--a systematic review. *Scand J Work Environ Health*. 2010;36(3):202–15.
81. Zusman EZ, Kapanen AI, Klaassen A, Reardon J. Workplace cardiovascular risk reduction by healthcare professionals—a systematic review. *Occup Med* (Chic Ill). 2021;71(6–7):270–6.
82. Rotunda W, Rains C, Jacobs SR, et al. Weight Loss in Short-Term Interventions for Physical Activity and Nutrition Among Adults With Overweight or Obesity: A Systematic Review and Meta-Analysis. *Prev Chronic Dis*. 2024;21:E21.
83. Lemstra M, Bird Y, Nwankwo C, Rogers M, Moraros J. Weight loss intervention adherence and factors promoting adherence: a meta-analysis. *Patient Prefer Adherence*. 2016;10:1547–59.
84. Tronieri JS, Wadden TA, Chao AM, Pearl RL, Alamuddin N, Berkowitz RI. Early Weight Loss in Behavioral Treatment Predicts Later Rate of Weight Loss and Response to Pharmacotherapy. *Annals of Behavioral Medicine*. 2019;53(3):290–5.
85. Fritsche A, Wagner R, Heni M, et al. Different Effects of Lifestyle Intervention in High- and Low-Risk Prediabetes: Results of the Randomized Controlled Prediabetes Lifestyle Intervention Study (PLIS). Diabetes. 2021;70(12):2785–95.
86. Rippe JM. Lifestyle Strategies for Risk Factor Reduction, Prevention, and Treatment of Cardiovascular Disease. *Am J Lifestyle Med*. 2019;13(2):204–12.
87. Leso V, Carugno M, Carrer P, et al. The Total Worker Health® (TWH) approach: a systematic review of its application in different occupational settings. *BMC Public Health*. 2024;24(1):2037.
88. European Commission. Publications Office of the European Union. 2022 (cited 2025 May 16). Healthier Together – EU Non-Communicable Diseases Initiative. Available from: https://health.ec.europa.eu/non-communicable-diseases/healthier-together_en
89. WHO. Implementation roadmap 2023–2030 for the Global action plan for the prevention and control of NCDs 2013–2030. 2025;
90. Galanti T, Guidetti G, Mazzei E, Zappalà S, Toscano F. Work From Home During the COVID-19 Outbreak: The Impact on Employees' Remote Work Productivity, Engagement, and Stress. *J Occup Environ Med*. 2021 1;63(7):e426–32.

La Medicina del Lavoro

Reviews, commentaries, perspectives

Med. Lav. 2025; 116 (6): 18179
DOI: 10.23749/mdl.v116i6.18179

Effectiveness of Occupational Health Promotion Programs on Cardiometabolic risk factors: A Systematic Review and Three-Level Meta-Analysis

SUPPLEMENTARY MATERIAL

Table S1. Search strings on different electronic databases: PubMed, Embase, Web of Science.

PubMed

#1	((("Health Promotion"[Mesh] OR "Health Promotion"[Title/Abstract] OR "Health Education"[Mesh] OR "Total Worker Health"[Title/Abstract] OR "Health Education*"[Title/Abstract] OR "Health Program*"[Title/Abstract] OR "Health improvement"[Title/Abstract] OR "Health Behavior"[Mesh] OR "Health Behavior"[Title/Abstract] OR "Health Campaign*"[Title/Abstract] OR (Wellness Program*[Title/Abstract]) OR "Wellbeing Program*"[Title/Abstract] OR "Health Enhancement"[Title/Abstract] OR "Risk Reduction Behavior"[Mesh] OR "Lifestyle Risk Reduction*"[Title/Abstract] OR "Healthy People Programs"[Mesh] OR "Health Incentive Program*"[Title/Abstract]) OR (Education Program*[Title/Abstract]))
#2	(("Work"[Mesh] OR "Workplace"[Mesh] OR "Workplace*"[Title/Abstract] OR "Job Site"[Title/Abstract] OR "Work-site"[Title/Abstract] OR "Work Site"[Title/Abstract] OR "Occupations"[Mesh] OR (Occupation*[Title/Abstract]) OR "Occupational Groups"[Mesh] OR (Occupational Groups*[Title/Abstract]) OR (Work Force*[Title/Abstract]) OR (Worker*[Title/Abstract]) OR (Work Staff*[Title/Abstract]) OR "Occupational Environment"[Title/Abstract] OR (Occupational Health Service*[Title/Abstract]) OR "Working Environment"[Title/Abstract] OR "Health Surveillance"[Title/Abstract]))
#3	("Blood Pressure"[Mesh] OR "Waist Circumference"[Mesh] OR (Waist Circumference[Title/Abstract]) OR "Body Mass Index"[Mesh] OR "Body Mass Index"[Title/Abstract] OR "Hematologic Tests"[Mesh] OR "Hematologic Tests"[Title/Abstract] OR (Blood Test*[Title/Abstract]) OR (Hematological Test*[Title/Abstract]) OR (Blood Analys*[Title/Abstract]) OR "Smoking Reduction"[Mesh] OR "Smoking Reduction"[Title/Abstract] OR "Smoking Cessation"[Mesh] OR "Smoking Cessation"[Title/Abstract] OR "Cholesterol*"[Mesh] OR "Body Weight"[Mesh] OR (Body Weight*[Title/Abstract]) OR "Triglycerides"[Mesh] OR (Triglycerid*[Title/Abstract]) OR "Weight Loss"[Mesh] OR "Weight Loss"[title/abstract] OR "Waist-Hip Ratio"[Mesh] OR "Diabetes Mellitus, Type 2/Prevention and Control"[Mesh] OR "Blood Glucose"[Mesh] OR "Blood Glucose"[title/abstract] OR "Hyperglycemia/Prevention and Control"[Mesh] OR "Insulin Resistance/Prevention and Control"[Mesh] OR "Cardiovascular Diseases/Prevention and Control"[Mesh] OR "Cholesterol/Blood"[Mesh] OR "Heart Diseases/Prevention and Control"[Mesh] OR "Exercise"[Mesh] OR "Exercise"[title/abstract] OR "Physical Activit*"[title/abstract]))
#4	**#1 AND #2 AND #3**

Embase

#1	'blood pressure':ti,ab OR 'waist circumference':ti,ab OR 'body mass':ti,ab OR 'blood examination':ti,ab OR 'blood analysis':ti,ab OR 'smoking reduction':ti,ab OR 'smoking cessation':ti,ab OR 'cholesterol':ti,ab OR 'body weight':ti,ab OR 'triacylglycerol':ti,ab OR 'body weight loss':ti,ab OR 'waist hip ratio':ti,ab OR 'diabetes mellitus':ti,ab OR 'glucose blood level':ti,ab OR 'hyperglycemia':ti,ab OR 'insulin resistance':ti,ab OR 'cardiovascular disease':ti,ab OR 'heart disease':ti,ab OR 'exercise':ti,ab OR 'physical activity':ti,ab'
#2	'work':ti,ab OR 'workplace':ti,ab OR 'occupation':ti,ab OR 'named groups by occupation':ti,ab OR 'workforce':ti,ab OR 'worker':ti,ab OR 'occupational health service':ti,ab OR 'work environment':ti,ab OR 'health survey':ti,ab
#3	'health promotion':ti,ab OR 'health education':ti,ab OR 'health program':ti,ab OR 'health behavior':ti,ab OR 'education program':ti,ab
#4	**#1 AND #2 AND #3**

Web of Science

#1	(TI=(health promotion)) OR AB=(health promotion)) OR TI=(health education)) OR AB=(health education)) OR TI=(total worker health)) OR AB=(total worker health)) OR TI=(health program)) OR AB=(health program)) OR TI=(Health Behavior)) OR AB=(Health Behavior)) OR TI=(Health Campaign)) OR AB=(Health Campaign)) OR TI=(Wellness Program)) OR AB=(Wellness Program)) OR TI=(Wellbeing Program)) OR AB=(Wellbeing Program)) OR TI=(Risk Reduction Behavior)) OR AB=(Risk Reduction Behavior)) OR TI=(Lifestyle Risk Reduction)) OR AB=(Lifestyle Risk Reduction)) OR TI=(Healthy People Programs)) OR AB=(Healthy People Programs)
#2	(TI=(work)) OR AB=(work)) OR TI=(workplace)) OR AB=(workplace)) OR TI=(occupation)) OR AB=(occupation)) OR TI=(occupational groups)) OR AB=(occupational groups)) OR TI=(job site)) OR AB=(job site)) OR TI=(work force)) OR AB=(work force)) OR TI=(worker)) OR AB=(worker)) OR TI=(work staff)) OR AB=(work staff)) OR TI=(Occupational Environment)) OR AB=(Occupational Environment)) OR TI=(Occupational Health Service)) OR AB=(Occupational Health Service)) OR TI=(Working Environment)) OR AB=(Working Environment)) OR TI=(Health Surveillance)) OR AB=(Health Surveillance))
#3	(TI=(Blood Pressure)) OR AB=(Blood Pressure)) OR TI=(Waist Circumference)) OR AB=(Waist Circumference)) OR TI=(Body Mass Index)) OR AB=(Body Mass Index)) OR TI=(Blood Tests)) OR AB=(Blood Tests)) OR TI=(Blood Analysis)) OR AB=(Blood Analysis)) OR TI=(Smoking Reduction)) OR AB=(Smoking Reduction)) OR TI=(Smoking Cessation)) OR AB=(Smoking Cessation)) OR TI=(Body Weight)) OR AB=(Body Weight)) OR TI=(Triglycerides)) OR AB=(Triglycerides)) OR TI=(Weight Loss)) OR AB=(Weight Loss)) OR TI=(Waist-Hip Ratio)) OR AB=(Waist-Hip Ratio)) OR TI=(Diabetes)) OR AB=(Diabetes)) OR TI=(Blood Glucose)) OR AB=(Blood Glucose)) OR TI=(Hyperglycemia)) OR AB=(Hyperglycemia)) OR TI=(Insulin Resistance)) OR AB=(Insulin Resistance)) OR TI=(Cardiovascular Diseases)) OR AB=(Cardiovascular Diseases)) OR TI=(Cholesterol)) OR AB=(Cholesterol)) OR TI=(Heart Diseases)) OR AB=(Heart Diseases)) OR TI=(Disease Management)) OR AB=(Disease Management)) OR TI=(Exercise)) OR AB=(Exercise)) OR TI=(Physical Activity)) OR AB=(Physical Activity))
#4	**#1 AND #2 AND #3**

Table S2. PICOS framework – inclusion and exclusion criteria.

	INCLUSION CRITERIA	EXCLUSION CRITERIA
P – POPULATION	Adult workers	Non-workers (general population, patients, etc.)
I – INTERVENTION	Health promotion interventions in workplaces related to cardiovascular risk factors	No intervention applied
C – COMPARISON	Workers who have not joined health promotion programs in the workplace	No comparison group
O – OUTCOME	Objective parameters related to cardiovascular risk factors	Health promotion programs targeting other health risks
S – STUDY DESIGN	Case control studies (RCT and Quasi-Experimental)	Other than case-control studies

Table S3. Study assumptions for SEs and Ess calculation.

Author, year	Outcome(s)	Assumption
Arrogi et al, 2019[23]	BMI, body fat %, WC	None
Asfar et al, 2021[24]	Smokers %	None
Barranco-Ruiz et al, 2019[25]	BMI, body weight, TOT-Chol, TG, SBP, DBP, body fat %, FBG, WHR	None
Clemes et al, 2022[26]	BMI, weight, TOT-Chol, HDL, LDL, TG, body fat %, WC, HbA1C	None
Day et al, 2019	Weight	None
Diaz-Benito et al, 2021[27]	BMI, weight, WC	r= 0.5 between baseline and follow-up samples in intervention and control groups
Fang et al, 2019[28]	BMI, weight, TOT-Chol, HDL, LDL, TG, SBP, DBP, body fat %, WC, FBG	r= 0.9 between baseline and follow-up samples in intervention and control groups
Garcia-Rojas et al, 2021[29]	SBP, DBP	None
Gerodimos et al, 2021[30]	SBP, DBP, body fat %	r= 0.5 between baseline and follow-up samples in intervention and control groups
Gimenez et al, 2024[31]	BMI, weight, TOT-Chol, HDL, LDL, TG, WC, FBG, HbA1C	r= 0.9 between baseline and follow-up samples in intervention and control groups
Guirado et al, 2024[32]	BMI, weight, TOT-Chol, HDL, LDL, TG, WC, FBG	None
Hassani et al, 2020[33]	BMI, weight, body fat %, FBG, HbA1C	None
Hee Woo et al, 2019[34]	BMI, TOT-Chol, HDL, LDL, TG, SBP, DBP, WC, FBG	r= 0.9 between baseline and follow-up samples in intervention and control groups
Hu et al, 2023[35]	SBP, DBP, smokers %	None
Iturriaga et al, 2019[36]	BMI, body fat %	r= 0.9 between baseline and follow-up samples in intervention and control groups
Jorvand et al, 2020[37]	TOT-Chol, HDL, LDL, TG, FBG	r= 0.5 between baseline and follow-up samples in intervention and control groups
Karatrantou et al, 2020[38]	SBP, DBP, body fat %	r= 0.5 between baseline and follow-up samples in intervention and control groups
Kim et al, 2023[39]	Smokers %	None
Kim et al, 2022[40]	TOT-Chol, HDL, LDL, TG, FBG	r= 0.9 between baseline and follow-up samples in intervention and control groups
Koch et al, 2022[41]	SBP, DBP	r= 0.9 between baseline and follow-up samples in intervention and control groups
Kong et al, 2022[42]	BMI, WC, WHR	None
Kotejoshyer et al, 2021[43]	BMI, body fat %	None
Kugathasan et al, 2023[44]	BMI, weight	r= 0.9 between baseline and follow-up samples in intervention and control groups
Lennefer et al, 2020[45]	BMI	r= 0.5 between baseline and follow-up samples in intervention and control groups
Ma et al, 2021[46]	BMI, body fat %	r= 0.9 between baseline and follow-up samples in intervention and control groups
Mahdavi-Roshan et al, 2020[47]	BMI, weight, WC	None

Maphong et al, 2021[48]	SBP, DBP, WC	r= 0.9 between baseline and follow-up samples in intervention and control groups
Mat Azmi et al, 2022[49]	TOT-Chol, HDL, LDL, TG, FBG	r= 0.9 between baseline and follow-up samples in intervention and control groups
Moon et al, 2024[50]	BMI, TOT-Chol, HDL, SBP, DBP, WC, FBG, HbA1C	r= 0.5 between baseline and follow-up samples in intervention and control groups
Nagata et al, 2022[51]	BMI, weight, LDL, SBP, DBP, WC, HbA1C	r= 0.9 between baseline and follow-up samples in intervention and control groups
Nahm et al, 2020[52]	BMI, weight, SBP, DBP, body fat %	r= 0.9 between baseline and follow-up samples in intervention and control groups
Ozaki et al, 2019[53]	BMI, weight	None
Raymond et al, 2019[54]	BMI, TOT-Chol, HDL, LDL, TG, WC, HbA1C	r= 0.5 between baseline and follow-up samples in intervention and control groups
Rigotti et al, 2020[55]	Smokers %	None
Röhling et al, 2020[56]	BMI, weight, TOT-Chol, HDL, LDL, TG, SBP, DBP, body fat %, WC, FBG, HbA1C	None
Ruettger et al, 2022[57]	BMI, weight, TOT-Chol, HDL, LDL, TG, SBP, DBP, body fat %, WC, HbA1C	None
Ryu et al, 2021[58]	BMI, TOT-Chol, HDL, LDL, TG, SBP, DBP, body fat %, WC	r= 0.9 between baseline and follow-up samples in intervention and control groups
Saavedra et al, 2020[59]	BMI, weight, TOT-Chol, HDL, LDL, TG, SBP, body fat %, WHR	r= 0.9 between baseline and follow-up samples in intervention and control groups
Shakerian et al, 2023[60]	BMI, body fat %, WHR	r= 0.9 between baseline and follow-up samples in intervention and control groups
Song et al, 2019[61]	BMI, TOT-Chol, HDL, SBP, DBP, FBG	None
Thorndike et al, 2021[62]	BMI, weight, TOT-Chol, HDL, LDL, TG, SBP, DBP, WC, HbA1C	None
Van de Ven et al, 2023[63]	BMI, weight, smokers %	r= 0.9 between baseline and follow-up samples in intervention and control groups
Wang et al, 2020[64]	SBP, DBP, smokers %	None
Wilson et al, 2022[65]	BMI, weight, SBP, DBP, body fat %, WC, WHR	r= 0.5 between baseline and follow-up samples in intervention and control groups

BMI = Body Mass Index; TOT-Chol = Total Cholesterol; HDL = High-Density Lipoprotein; LDL = Low-Density Lipoprotein; TG = Triglycerides; SBP = Systolic Blood Pressure; DBP = Diastolic Blood Pressure; WC = Waist Circumference; FBG = Fasting Blood Glucose; HbA1C = Glycated Haemoglobin; WHR = Waist-to-Hip Ratio.

Table S4. Quality assessment criteria.

Criterion	Range	Score	Description
Design	0-1	1	if randomised trial
		0	if quasi-experimental design of any kind
Assessment of intervention	0-1	1	if the intervention has been clearly defined and measured
		0	if the intervention has not been clearly defined and measured
Assessment of outcome	0-1	1	if the outcome has been clearly defined and measured
		0	if the outcome has not been clearly defined and measured
Control for confounding	0-1	1	if RCT or sufficient/appropriate control for major confounders
		0	if insufficient control for major confounders
Evidence of selection bias	0-1	1	if absence of evidence for selection bias
		0	If substantial presence of evidence for selection bias

Table S5. Numbers of studies investigating different outcomes.

Outcome	**N. of studies**
BMI (kg/m2)	30
Weight (kg)	19
Total Cholesterol (mg/dL)	16
HDL Cholesterol (mg/dL)	16
LDL Cholesterol (mg/dL)	15
Triglycerides (mg/dL)	11
Systolic Blood Pressure (mmHg)	20
Diastolic Blood pressure (mmHg)	19
Body fat (%)	16
Waist circumference (cm)	17
Glucose (mmol/L)	11
Smoking (%)	5

Table S6. Quality assessment of included studies.

FIRST AUTHOR	YEAR	QUALITY ASSESSMENT						
		A	B	C	D	E	TOT	SCORE
Arrogi et al	2019	0	1	1	0	1	3	Low
Asfar et al	2021	1	1	1	1	1	5	High
Barranco-Ruiz et al	2019	1	1	1	1	0	4	High
Clemes et al	2022	1	1	1	1	1	5	High
Day et al	2019	1	1	1	1	0	4	High
Diaz-Benito et al	2021	1	1	1	1	0	4	High
Fang et al	2019	1	1	1	1	1	5	High
Garcia-Rojas et al	2021	0	1	1	1	0	3	Low
Gerodimos et al	2021	1	1	1	1	1	5	High
Gimenez et al	2024	1	1	1	1	0	4	High
Guirado et al	2024	1	1	1	1	0	4	High
Hassani et al	2020	1	1	1	1	1	5	High
Hee Woo et al	2019	1	1	1	1	0	4	High
Hu et al	2023	1	1	1	1	1	5	High
Iturriaga et al	2019	1	1	1	1	0	4	High
Jorvand et al	2020	1	1	1	0	0	3	Low
Karatrantou et al	2020	1	1	1	1	0	4	High
Kim et al	2023	0	1	1	1	1	4	Low
Kim et al	2022	1	1	1	1	0	4	High
Koch et al	2022	1	1	1	1	1	5	High
Kong et al	2022	1	1	1	1	1	5	High
Kotejoshyer et al	2021	1	1	1	1	1	5	High
Kugathasan et al	2023	0	1	1	1	0	3	Low
Lennefer et al	2020	1	1	1	1	0	4	High
Ma et al	2021	1	1	1	1	0	4	High
Mahdavi-Roshan et al	2020	0	1	1	1	1	4	High
Maphong et al	2021	0	1	1	0	0	2	Low
Mat Azmi et al	2022	0	1	1	0	0	2	Low
Moon et al	2024	0	1	1	0	0	2	Low
Nagata et al	2022	0	1	1	0	0	2	Low
Nahm et al	2020	1	1	1	1	0	4	High
Ozaki et al	2019	1	1	1	1	1	5	High
Raymond et al	2019	0	1	1	0	0	2	Low
Rigotti et al	2020	1	1	1	1	1	5	High
Röhling et al	2020	1	1	1	1	1	5	High
Ruettger et al	2022	1	1	1	1	1	5	High
Ryu et al	2021	0	1	1	1	1	4	High
Saavedra et al	2020	0	1	1	0	0	2	Low
Shakerian et al	2023	1	1	1	1	0	4	High
Song et al	2019	1	1	1	1	1	5	High
Thorndike et al	2021	1	1	1	1	1	5	High
Van de Ven et al	2023	1	1	1	1	0	4	High
Wang et al	2020	1	1	1	1	1	5	High
Wilson et al	2022	0	1	1	0	0	2	Low

Table S7. Aggregate characteristics of included studies.

Publication details		Workers' details		Workplace characteristics		Intervention characteristics	
Geographic area		**Total participants**	**49813**	**Setting**		**Duration (months)**	9.8 ± 13.4
Asia	17 (39%)	N cases	32457 (%)	Hospital	9 (20%)	**Area of intervention**	
Europe	15 (34%)	N controls	17356 (	Industry	9 (20%)	Physical activity	38/44 (86%)
North America	8 (18%)	Median sample size	110 (16-24396)	Tertiary	16 (36%)	Dietary habits	21/44 (48%)
Other	4 (9%)	Mean age	41.0 ± 5.4	Mixed/unspecified	10 (23%)	Smoking reduction	8/44 (18%)
Year of publication		**Gender (male%)**	54.5 ± 33.6	**Number of sites**	6.6 (±9.7)	Stress management	6/44 (14%)
≥ 2022	28 (64%)	**Designation**		**Company dimension**		Alcohol drinking	2/44 (5%)
< 2022	16 (36%)	White collar	32 (73%)	Large	24 (54%)	Sleep hygiene	1/44 (2%)
Study design		Blue collar	4 (9%)	Medium	6 (14%)	**Modality of intervention**	
RCT	30 (68%)	Mix/unspecified	8 (18%)	Unspecified	14 (32%)	In-person	19/44 (43%)
Quasi-experimental	14 (32%)	**Health status**				Web-Online	4/44 (9%)
Randomization (for 30 RCTs)		Healthy	18 (41%)			Mixed	21/44 (48%)
Cluster	7 (23%)	CV risk factors	11 (25%)			**Single vs multiple areas of intervention**	
Individual	23 (77%)	Mixed	15 (34%)			Multiple	19 (43%)
Study quality						Single	25 (57%)
Low	27 (61%)					**Main author of the intervention**	
High	17 (39%)					Physician	15 (34%)
						Other sanitary	2 (5%)
						Other	27 (61%)
						N intervention/months	14.2 (13.7)
						Mangement involvement	
						Yes	10 (23%)
						No	34 (77%)
						Economic incentives	
						Yes	10 (23%)
						No	34 (77%)

Figure S1. Forest plot of weight.

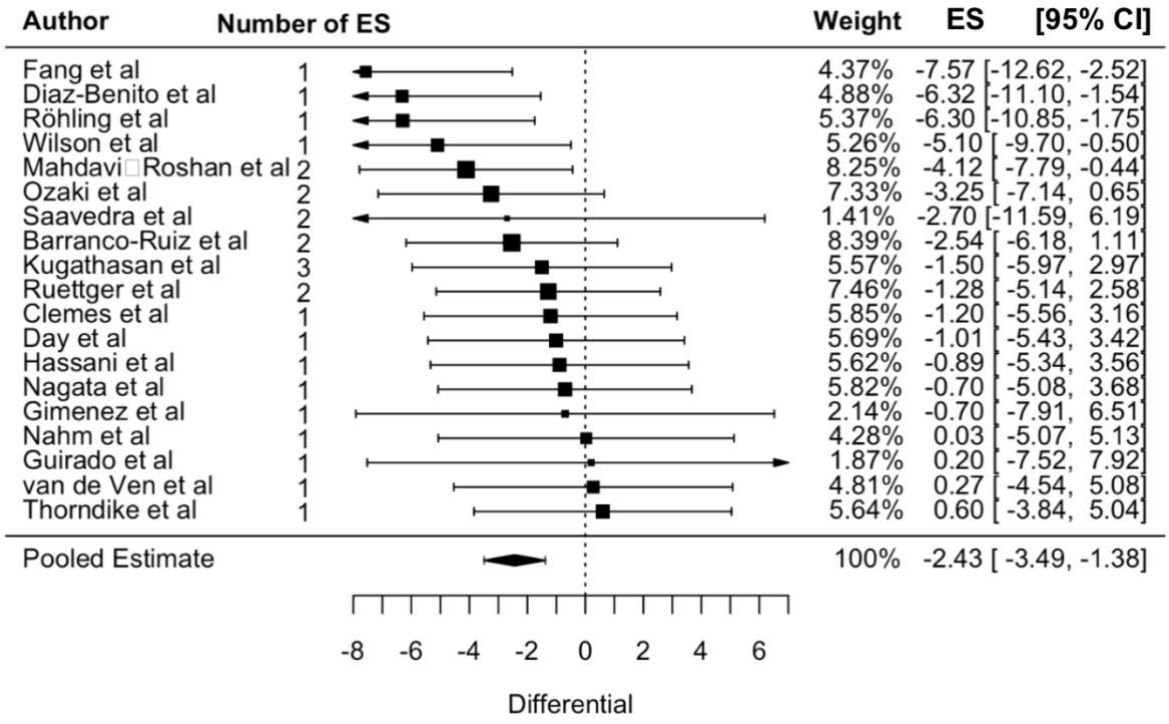

Figure S2. Forest plot of body fat.

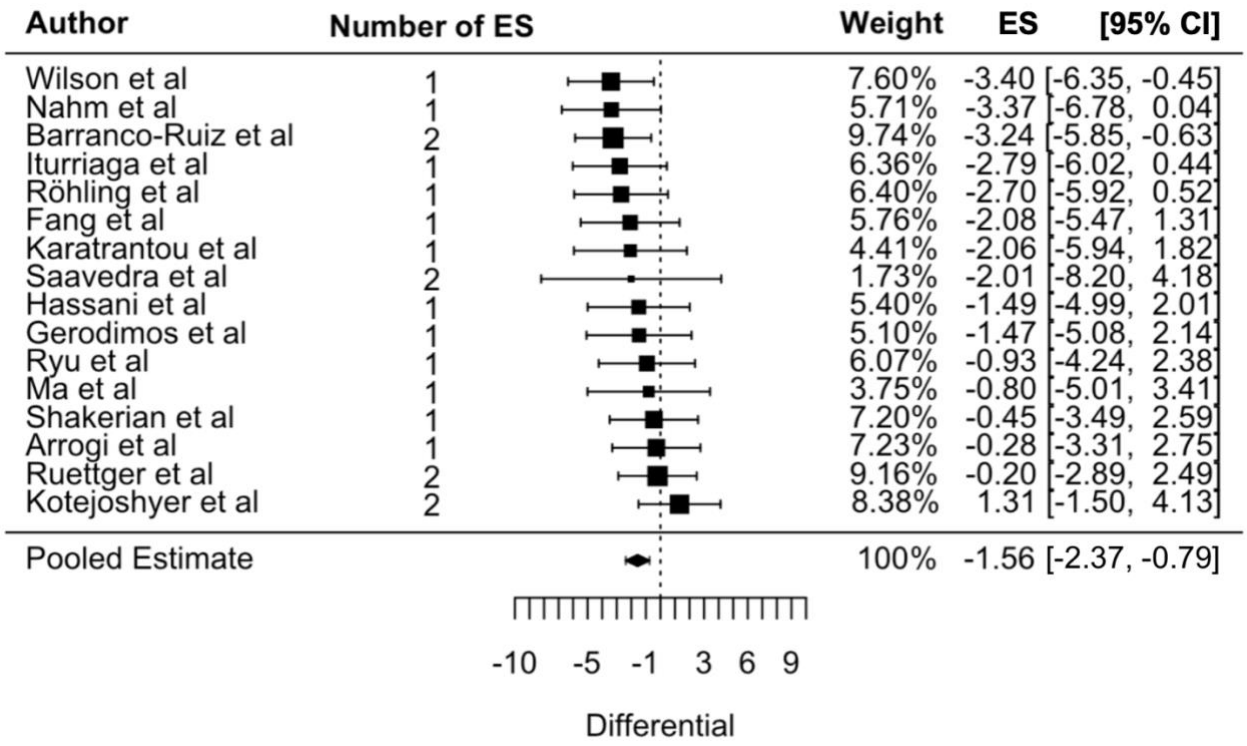

Figure S3. Forest plot of waist circumference.

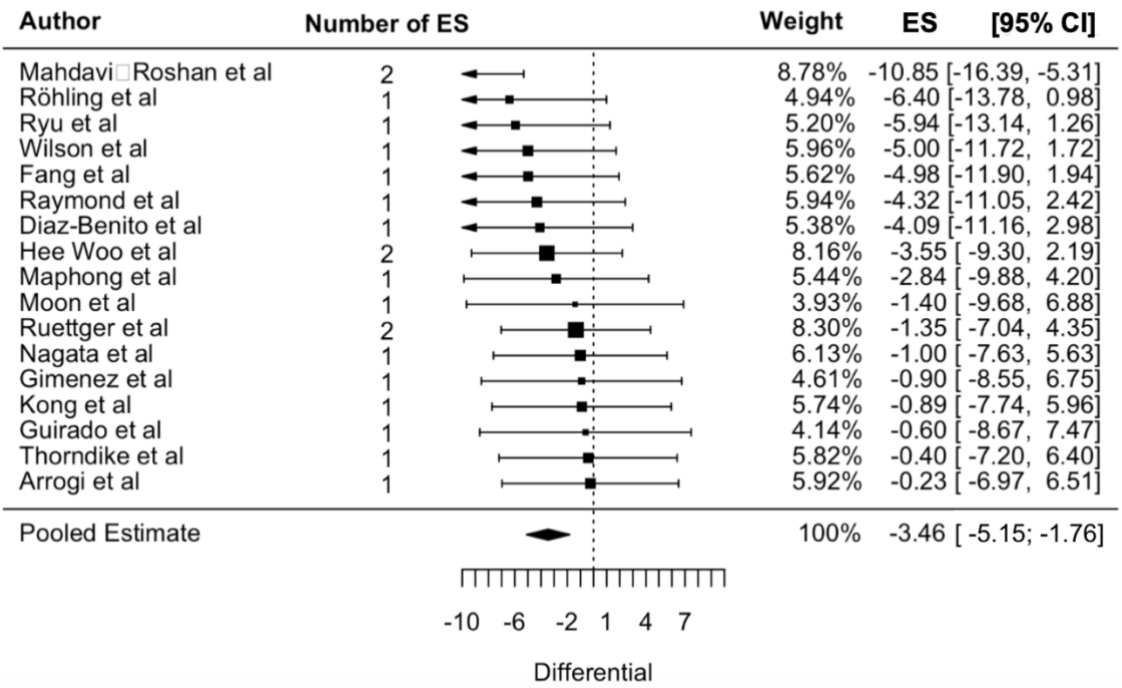

Figure S4. Forest plot of total cholesterol.

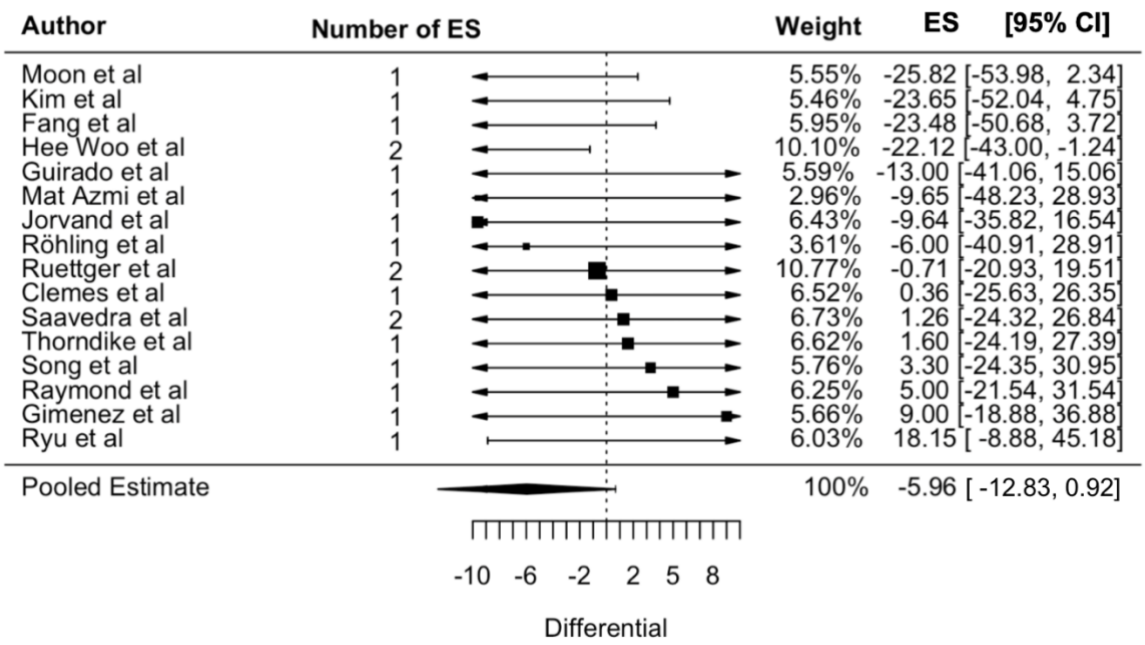

Figure S5. Forest plot of HDL cholesterol.

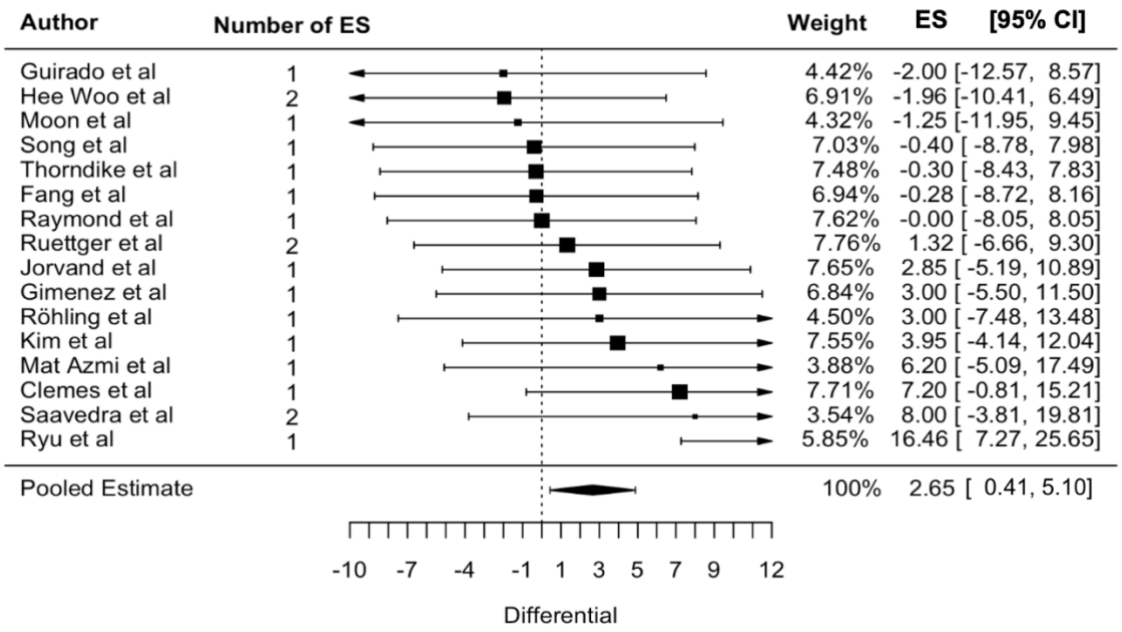

Figure S6. Forest plot of LDL cholesterol.

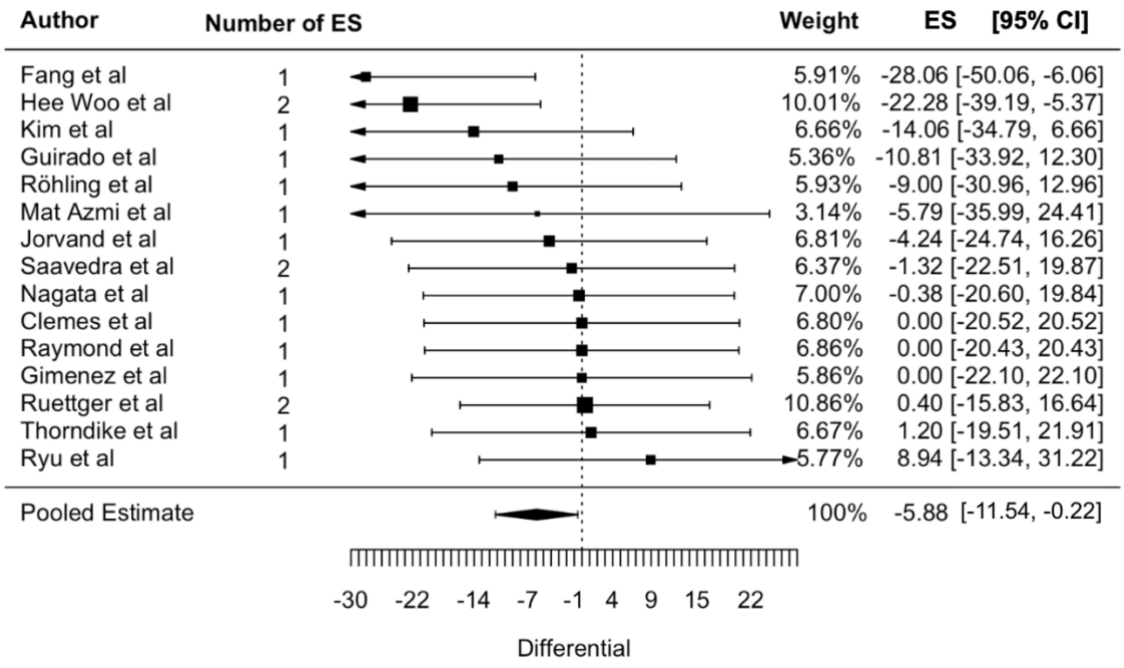

Figure S7. Forest plot of DBP.

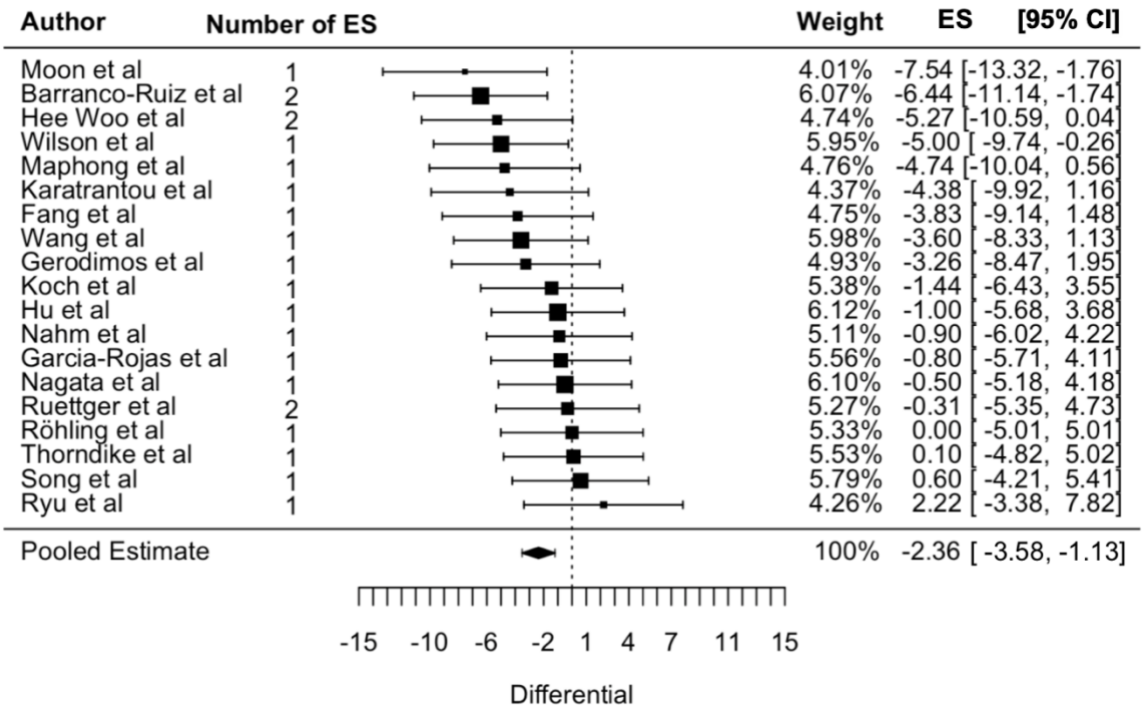

Figure S8. Forest plot of FBG.

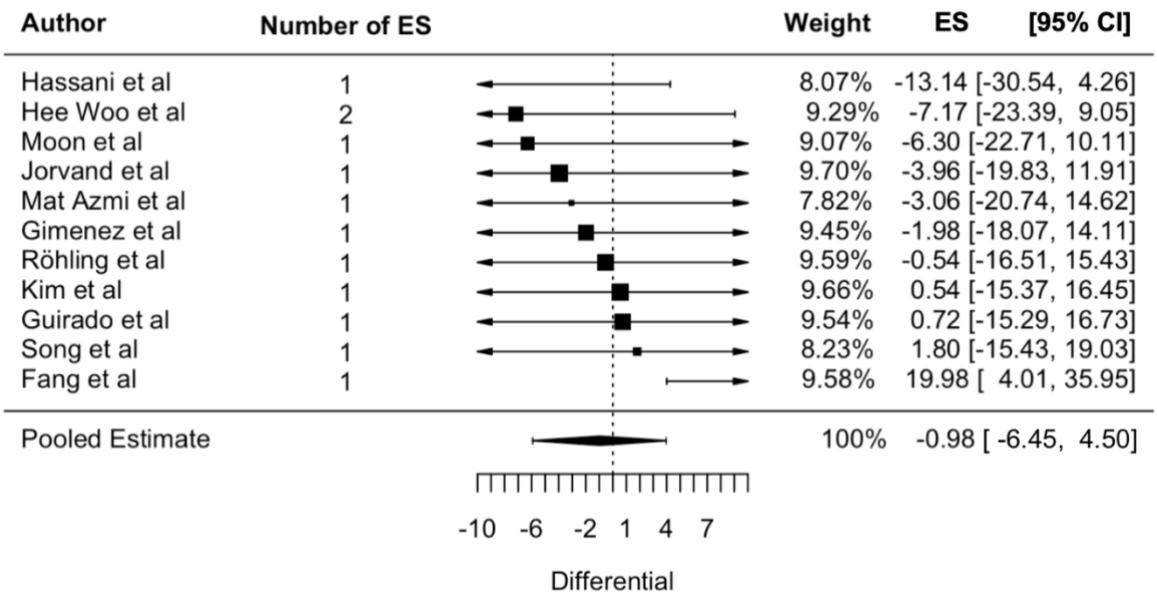

Figure S9. Forest plot of triglycerides.

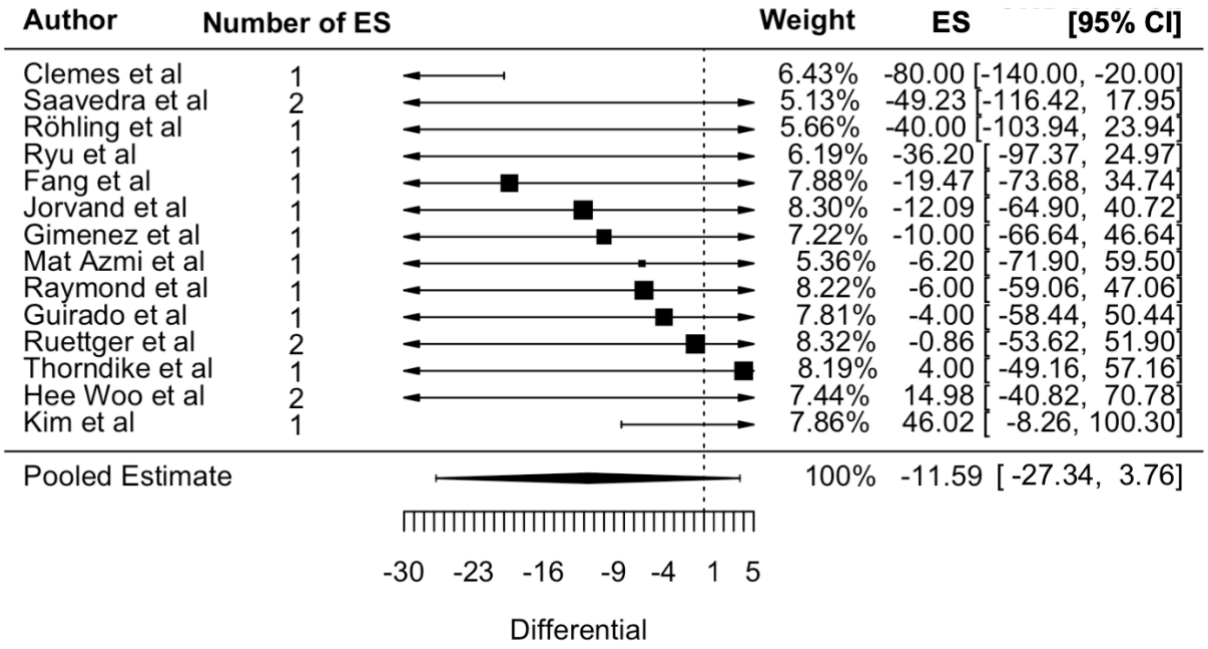

Figure S10. Forest plot of smoking habit.

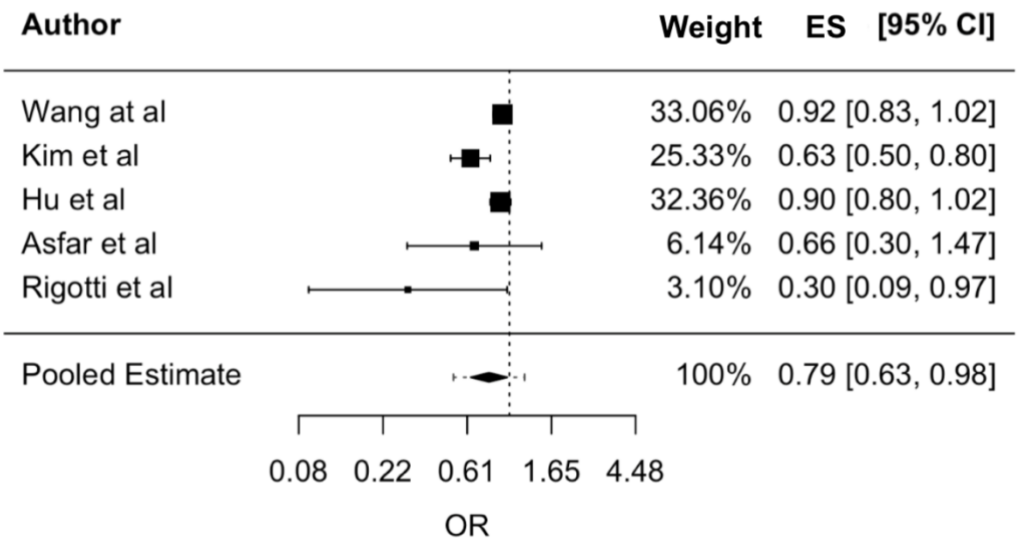

Table S8. Stratified meta-analyses and univariate meta-regressions results from three-levels random effects models.

	BMI (kg/m²)		Weight (kg/m²)		Body fat (%)		Total Cholesterol (mg/dL)		LDL Cholesterol (mg/dL)		HDL Cholesterol (mg/dL)	
	n ES	β (95% CI)	n ES	β (95% CI)	n ES	β (95% CI)	n ES	β (95% CI)	n ES	β (95% CI)	n ES	β (95% CI)
Overall estimate	39	**-0.63 (-0.92; -0.34)**	26	**-2.44 (-3.42; -1.44)**	20	**-1.58 (-2.37; -0.79)**	19	-5.96 (-12.84; 0.92)	**18**	**-5.88 (-11.54; -0.21)**	19	**2.75 (0.41; 5.10)**
Study Design												
RCT	26	-0.49 (-0.81; -0.18)	17	-2.60 (-3.49; -1.02)	26	-1.53 (-2.48; -0.58)	12	-7.88 (-15.73; -0.04)	11	-9.05 (-15.76; -2.33)	12	1.61 (-1.15; 4.38)
Quasi-experimental	13	-0.92 (-1.52; -0.33)	9	-2.84 (-4.84; -0.85)	13	-1.66 (-3.37; 0.06)	7	-2.46 (-14.26; 9.35)	7	-0.17 (-9.19; 8.85)	7	5.05 (1.11; 9.00)
Test of group difference (p.value)		0.214		0.610		0.869		0.461		0.113		0.150
Geographic region												
North America	8	0.02 (-0.59; 0.62)	5	-0.69 (-3.20; 1.81)	2	1.33 (-0.32; 2.99)	6	3.25 (-10.94; 17.46)	2	0.59 (-14.85; 16.03)	3	-0.23 (-0.40; 8.31)
Asia	15	-0.95 (-1.36; -0.49)	8	-2.79 (4.76; -1.00)	6	**-1.47 (-2.52; -0.42)**	7	**-**15.28 (-24.72; -5.50)	7	-11.53 (-20.06; -3.00)	7	3.42 (-0.54; 7.39)
Europe	12	-0.40 (-0.98; 0.07)	9	-2.62 (-4.50; -0.74)	9	**-1.27 (-2.16; -0.38)**	8	-3.03 (-13.04; 6.98)	8	-3.18 (-2.694; 0.625)	8	3.93 (-0.40; 8.31)
Other	4	-1.14 (-1.99; -0.28)	4	-3.14 (-5.85; -0.43)	3	**-3.30 (-4.48; -2.11)**	1	9.00 (-1.72; 19.72)	1	0.000 (-8.96; 8.96)	1	3.00 (-6.52; 12.52)
Test of group difference (p.value)		0.051		0.458		**0.001**		0.086		0.352		0.602
Workplace setting												
Industry	8	-0.45 (-1.05; 0.16)	3	-2.66 (-5.45; 0.12)	6	-1.25 (-2.28; -0.27)	2	-10.21 (-36.45; 16.03)	1	-28.06 (-48.80; -7.32)	2	-0.34 (-2.236; 1.549)
Healthcare	7	-1.19 (-2.17; -0.24)	5	-3.07 (- 5.35; -0.80)	2	-2.33 (-3.51; -1.15)	6	-8.26 (-21.29; 4.77)	6	-9.298 (-17.18; -0.71)	6	1.41 (-0.542; 3.175)
Tertiary	17	-0.53 (-0.93; -0.14)	14	-1.82 (-3.31; -0.31)	11	-1.24 (-2.52; 0.03)	9	-2.73 (-11.42; 5.95)	9	-0.56 (-7.86; 6.73)	9	5.41 (0.733; 8.848)
Various	3	-0.643 (-1.25; -0.03)	3	-2.13 (-4.87; 1.09)	0		2	-8.86 (-36.92; 19.20)	2	-6.920 (-20.70; 6.68)	2	1.96 (-1.917; 5.829)
Test of group difference (p.value)		0.603		0.780		0.840		0.867		0.079		0.296
Dimension ef enterprise												
Large	23	-0.65 (-1.05; -0.24)	12	-2.27 (-3.83; -0.71)	13	-1.42 (-2.50; -0.32)	11	-7.50 (-15.61; 1.36)	10	-5.40 (-12.67; 2.10)	11	0.79 (-0.48, 2.06)
Medium	6	-0.34 (-0.68; 0.00)	4	-3.30 (-6.66; 0.06)	4	-1.37 (-3.74; 0.98)	3	9.48 (-9.14; 28.10)	3	3.67 (-12.78; 20.13)	3	**13.83 (9.01; 18.66)**
Test of group difference (p.value)		0.257		0.562		0.973		0.100		0.292		**0.001**
Job designation												
Blue collar	5	-0.37 (-1.17;0.43)	4	-1.17 (-3.29; 0.94)	4	-0.54 (-2.24; 1.15)	3	-0.15 (-2.74; 2.44)	3	0.255 (-13.26; 13.78),	3	4.27 (-2.34; 10.89)
White collar	30	-0.74 (-1.11; -0.38)	21	-2.89 (- 3.94; -1.85)	16	-1.84 (-2.72; -0.94)	14	-9.23 (-17.76; -0.69)	14	-7.92 (-14.52; -1.32)	14	3.07 (0.02; 6.12)
Mixed	3	-0.00 (-1.02; 1.00)	1	0.270 (-4.27; 4.81)	0		1	3.30 (-6.80; 13.40)	0		1	-0.40 (10.15; 9.35)
Test of group difference (p.value)		0.314		0.165		0.174		0.483		0.265		0.702

Only pathological												
Yes	12	**-1.45 (-2.00; -0.90)**		-3.73 (-5.47; -1.98)	7	-1.69 (-3.75; 0.37)	3	**-22.52 (-36.59; -8.63)**	6	**-24.17 (-33.09; -15.24)**	5	-0.58 (6.79; 5.63)
No	24	**-0.38 (-0.67; -0.08)**		-1.19 (-3.03; -0.90)	9	-1.53 (-2.44; -0.62)	16	-2.52 (-8.93; 3.80)	10	-2.16 (-6.11; 1.78)	12	3.29 (0.79; 5.78)
Test of group difference (p.value)		**0.001**		0.073		0.885		**0.014**		**0.001**		0.239
Modality of intervention												
In person	13	-0.61 (-1.16; -0.05)	7	-2.89 (-5.21; -0.57)	13	-1.45 (-2.54; -0.36)	5	-6.86 (-14.85; 1.13)	5	-9.151 (-21.023; 2.270)	5	0.07 (-4.17; 6.18)
Web	1	-0.23 (-1.01; 0.55)	1	-1.01 (-5.58; 3.57)	0		2	-15.52 (-29.07; -1.97)	2	-9.029 (-18.653; 0.595)	2	3.40 (-3.04; 9.84)
Mixed	25	-0.65 (-1.01; -0.29)	18	-2.39 (-3.65; -1.12)	7	-1.522 (-3.01; -0.42)	11	-4.08 (-12.80; 4.63)	11	-4.223 (-11.057; 2.611)	12	3.28 (0.12; 6.44)
Test of group difference (p.value)		0.625		0.747		0.742		0.496		0.643		0.716
Author of intervention												
Physician	13	**-0.75 (-1.25; -0.25)**	10	-3.03 (-4.90; -1.15)	4	-2.98 (-4.03; -0.64)	9	-11.04 (-21.05; -1.03)	8	-10.220 (-18.39; -1.80)	9	0.84 (-2.46; 4.14)
Other sanitary	3	**-1.60 (-2.55; -0.65)**	3	-2.76 (-5.75; 0.23)	1	-1.49 (-5.03; 2.05)	0		0		0	
Other	23	**-0.39 (-0.76; -0.02)**	13	-1.94 (-3.44; -0.45)	15	-1.24 (-2.25; -0.41)	10	-1.20 (-8.60; 6.21)	10	-2.645 (-10.06; 4.34)	10	4.28 (1.29; 7.27)
Test of group difference (p.value)		**0.049**		0.632		0.456		0.121		0.181		0.121
Economic incentives												
Yes	11	-0.21 (-0.76; 0.34)	6	-0.50 (-2.36; 1.34)	0		5	-6.64 (-21.82; 8.54)	5	-7.24 (-17.81; 3.33)	5	-0.38 (-4.26; 3.50)
No	28	-0.79 (-1.13; -0.42)	20	**-2.99 (-3.99; -2.00)**	20		14	-5.85 (-21.82; 8.54)	13	-5.063 (-12.30; 1.68)	14	4.03 (1.50; 6.55)
Test of group difference (p.value)		0.086		**0.002**				0.927		0.751		0.060
Planning involvement of management												
Yes	10	-0.35 (-0.95; 0.26)	4	-0.90 (-3.98; 2.10)	4	-1.29 (-2.73; 0.16)	4	-2.15 (-21.99; 17.69)	3	2.14 (-11.90; 16.19)	4	5.30 (-2.81; 13.58)
No	29	-0.70 (-1.05; -0.35)	22	-2.64 (-3.69; -1.53	16	-1.48 (-2.41; -0.56)	15	-7.22 (-13.85; -0.59)	15	-7.46 (-13.21; -1.71)	15	2.06 (0.48; 3.47)
Test of group difference (p.value)		0.315		0.300		0.740		0.635		0.205		0.246
Study quality												
Low	13	-0.921 (-1.52; -0.36)	10	-2.26 (-3.49; 1.03)	14	-1.66 (-3.30; -0.03)	7	-2.46 (-14.26; 9.35)	9	-5.696 (-14.590; 3.199)	9	5.06 (1.11; 9.00)
High	26	-0.49 (-0.81; -0.18)	16	-2.84 (-4.84; -0.85)	6	-1.40 (-2.31; -0.48)	12	-7.88 (-15.73; -0.04)	9	-6,206 (-12.520; 0.108)	10	1.61 (-1.15; 4.38)
Test of group difference (p.value)		0.241		0.610		0.896		0.453		0.927		0.150
Duration of intervention (cat)												
< 3 months	21	**-0.93 (-1.33; -0.52)**	14	-3.35 (-4.63; -2.06)	10	**-2.14 (-3.08; -1.00)**	11	-8.77 (-19.54; 1.68)	11	-8.930 (-17.267; -0.593)	11	3.47 (-0.15; 7.09)
4 - 12 months	9	-0.57 (-1.11; -0.02)	7	-1.97 (-3.55; -0.40)	8	**-1.62 (-2.63; -0.62)**	4	-8.912 (-17.06; 1.76)	4	-4.294 (-10.904; 2.315)	4	1.68 (-2.62; 8.04)
> 12 months	9	-0.04 (-0.61; 0.54)	5	-0.85 (-2.96; 1.25)	2	1.39 (-0.92; 4.00)	4	2.51 (0.193; 3.107)	3	0.249 (-1.825; 2.234)	4	2.71 (-2.99; 6.35)

Test of group difference (p.value)		**0.047**		0.105		**0.034**		**0.027**		0.435		0.858
Mean age (p.value)	33	-0.03 (0.408)	23	0.03 (0.787)	17	-0.07 (0.410)	16	0.22 (0.735)	15	-0.16 (0.691)	16	-0.138 (0.582)
N intervention/month (p.value)	39	-0.01 (0.197)	26	-0.10 (0.846)	20	-0.02 (0.930)	19	0.00 (0.981)	18	0.17 (0.359)	19	**0.122 (0.025)**
Males (%) (p.value)	35	0.01 (0.177)	25	-0.006 (0.667)	19	0.02 (0.106)	19	0.03 (0.706)	16	0.06 (0.216)	17	-0.006 (0.837)
Study size (p.value)	37	0.00 (0.234)	24	0.00 (0.269)	19	0.01 (0.072)	19	0.01 (0.256)	18	0.002 (0.427)	19	-0.002 (0.247)

**Results in bold indicate statistical significance at 5% level.*

Table S8. continued

	Waist Circumference (cm)		**Glucose (mg/dL)**		**Diastolic blood pressure (mmHg)**		**Systolic blood pressure (mmHg)**		**Triglycerides**	
	n ES	**β (95% CI)**	**n ES**	**β (95% CI)**	**n ES**	**β (95% CI)**	**n ES**	**β (95% CI)**	**n ES**	**β (95% CI)**
Overall estimate	20	**-3.45 (-5.15; -1.76)**	12	-1.56 (-6.28; 3.17)	22	**-2.36 (-3.58; -1.13)**	24	**-3.75 (-5.68; -1.81)**	17	-11.78 (-27.32; 3.74)
Study Design										
RCT	11	-2.57 (-4.87; -0.28)	11	0.29 (-6.27; 6.85)	16	-2.23 (-3.75; -0.72)	16	-3.02 (-5.41; -0.63)	11	-5.19 (-23.07; 12.70)
Quasi-experimental	9	-4.50 (-6.93; -2.07)	3	-4.49 (-15.39; -6.40)	6	-2.63 (-4.91; -0.36)	8	-5.12 (-8.44; -1.81)	6	-11.738 (-18.49; -4.98)
Test of group difference (p.value)		0.240		0.422		0.764		0.298		0.502
Geographic region										
North America	2	-2.37 (-6.28; 1.451)	1	1.80 (-5.04; 8.64)	2	0.35 (-3.09; 3.80)	2	0.95 (-0.67; 2.57)	2	-1.14 (-10.94; 8.65)
Asia	10	-4.42 (-7.62; -2.86)	7	-2.43 (-10.59; 5.73)	10	-2.63 (-4.36; -0.91)	10	-3.87 (-6.49; -0.91)	6	0.16 (-21.56; 25.61)
Europe	6	-2.39 (-4.14; -0.27)	3	-0.12 (-1.74; 1.50)	6	-1.74 (-4.07; 0.58)	8	-3.49 (-5.72; -0.41)	8	-28.13 (-47.62; -2.58)
Other	2	-3.19 (-8.70; 2.30)	1	-1.98 (-4.93; 0.97)	4	-4.19 (-6.94; -1.43)	4	-6.63 (-10.75; -3.43)	1	-10.00 (-30.91; 10.91)
Test of group difference (p.value)		0.731		0.608		0.203		0.231		0.387
Workplace setting										
Industry	2	-2.55 (-7.19; 2.12)	3	3.72 (-15.76; 22.03)	3	-1.24 (-4.51; 2.01)	3	-1.40 (-3.57; 0.87)	1	-19.47 (-32.45; -6.49)
Healthcare	7	-5.06 (-8.07; -2.97)	5	-3.39 (-5.99; -1.14)	6	-1.73 (-4.26; 0.79)	6	-1.83 (-4.26; 0.73)	6	-2.90 (-14.633; 8.82)
Tertiary	15	-2.57 (-5.54; 0.39)	3	-2.82 (-7.31; 2.24)	11	-3.11 (-5.01; -1.21)	13	-5.52 (-8.03; -2.51)	8	-24.67 (-46.81; -2.53)
Various	1	-4.32 (-12.02; 3.39)	1	0.54 (-1.18; 2.26)	2	-2.29 (-4.82; 0.27)	2	-3.23 (-8.23; 1.76)	2	19.68 (-31.29; 70.66)
Test of group difference (p.value)		0.645		0.753		0.691		0.300		0.170
Dimension ef enterprise										
Large	14	-3.64 (-5.90; -1.38)	7	-3.81 (-6.24; -1.39)	18	-2.47 (-3.86; -1.07)	18	-4.02 (-6.24; -1.77)	9	-3.26 (-11.05; 4.52)
Medium	3	-3.56 (-8.20; 1.07)	1	-13.14 (-20.41; -5.87)	2	0.19 (-3.84; 4.21)	4	-1.47 (-6.79; 3.85)	3	**-42.29 (-70.98 -13.62)**

Test of group difference (p.value)		0.975		0.078		0.207		0.369		**0.015**
Job designation										
Blue collar	3	-0.88 (-5.43; 3.66)	1	-13.14 (-20.41; -5.87)	2	-0.33 (-2.48; 1.82)	2	-0.23 (-3.33; 2.91)	3	-25.30 (-75.24; 24.63)
White collar	16	-3.86 (-5.78; -1.93)	10	-0.84 (-5.97; 4.29)	16	-3.23 (-4.63; -1.84)	18	-4.55 (-6.73; -2.38)	13	-6.59 (-20.96; 7.79)
Mixed	0		1	1.80 (-5.04; 8.64)	3	-1.35 (-3.73; 1.01)	3	-2.11 (-5.86; 1.57)	0	
Test of group difference (p.value)		0.221		0.377		0.463		0.449		0.408
Only pathological										
Yes	8	**-6.14 (-8.92; -3.35)**	4	0.64 (-10.50; 11.78)	6	-4.10 (-6.41; -1.79)	6	-5.80 (-9.55; -2.05)	3	-0.58 (-6.79; 5.63)
No	12	**-2.45 (-4.26; -0.65)**	8	-1.59 (-8.36; 5.17)	16	-1.76 (-3.09; -0.43)	16	-3.02 (-5.24; -0.79)	14	3.29 (0.79; 5.78)
Test of group difference (p.value)		**0.03**		0.710		0.083		0.199		0.239
Modality of intervention										
In person	3	-3.56 (-7.90; 0.84)	3	3.33 (-7.97; 14.64)	9	-4.09 (-7.00; -1.18)	11	-4.46 (-6.63; -2.28)	5	-11.76 (-21.48; -2.04)
Web	0		2	-1.71 (-15.14; 11.71)	0		0		2	16.612 (-40.330; 73.555)
Mixed	17	-3.43 (-5.34; -1.52)	7	-3.44 (-10.98; 5.13)	13	-3.44 (-6.16; -0.72)	13	-3.133 (-5.850; -0.416)	10	-11.935 (-29.291; 5.421)
Test of group difference (p.value)		0.952		0.605		0.304		0.738		0.626
Author of intervention										
Physician	8	**-3.58 (-4.858; -2.321)**	5	0.77 (-11.79; 9.61)	8	-4.21 (-5.33; -1.88)	10	-2.88 (-5.34; -0.43)	8	-5.94 (-2.46; 4.54)
Other sanitary	2	**-10.85 (-14.19; -7.50)**	1	-13.14 (-20.41; -5.87)	0		0		0	
Other	10	**-2.04 (-3.72; -0.36)**	6	-0.39 (-1.14; 1.02)	14	-3.07 (-3.44; -0.53)	14	-4.35 (-6.85; -1.86)	9	4.97 (1.72; 8.22)
Test of group difference (p.value)		**0.001**		0.376		0.560		0.410		0.099
Economic incentives										
Yes	6	-2.187 (-5.14; 0.84)	3	-4.86 (-11.09; 1.37)	5	-1.07 (-3.65; 1.50)	5	-1.08 (-3.85; 0.74)	4	2.73 (-6.81; 2.26)
No	14	-4.223 (-6.38; -2.06)	9	-0.59 (-6.43; 5.25)	17	-2.72 (-4.09; -1.34)	19	**-4.76 (-6.77; -2.76)**	13	-16.59 (-33.85; 0.65)
Test of group difference (p.value)		0.278		0.737		0.832		**0.034**		0.055
Planning involvement of management										
Yes	5	-2.97 (-5.09; -0.85)	2	-5.53 (-9.38; -1.69)	4	-2.64 (-5.59; 0.30)	4	-6.79 (-11.18; -2.39)	3	-12.77 (-30.78; 5.23)
No	12	-3.81 (-5.83; -1.79)	10	-0.99 (-6.52; 4.53)	18	-2.29 (-3.69; -0.89)	20	-3.06 (-5.13; -0.99)	14	-9.53 (-25.45; 6.39)
Test of group difference (p.value)		0.736		0.516		0.254		0.125		0.791
Study quality										
Low	11	-4.51 (-6.93; -2.07)	2	-4.14 (-5.43; -2.84)	8	-2.67 (-4.91; -0.35)	8	-5.12 (-8.44; -1.81)	6	-11.74 (-18.49; -4.98)
High	9	-2.57 (-4.86; -0.28)	9	1.477 (-3.930; 6.883)	14	-2.23 (-3.75; -0.72)	16	-3.02 (-5.41; -0.63	11	-5.19 (-23.07; 12.70)

Test of group difference (p.value)		0.241		0.422		0.746		0.297		0.502
Duration of intervention (cat)										
< 3 months	13	-4.38 (-6.52; -2.23)	9	-1.15 (-8.27; 4.36)	10	-2.50 (-4.37; -0.92)	12	-3.89 (-6.73; -1.04)	10	-12.67 (-25.64; 0.31)
4 - 12 months	5	-2.37 (-5.14; 1.49)	2	-1.73 (-6.15; 2.67)	8	-3.32 (-5.26; -0.95)	8	-5.26 (-8.67; -1.85)	4	7.55 (-17.03; 32.13)
> 12 months	2	-1.82 (-6.23; 1.45)	1	1.80 (-5.04; 8.64)	4	-1.03 (-3.57; 1.55)	4	-1.24 (5.34; 2.86)	3	-25.59 (-75.74; 24.56)
Test of group difference (p.value)		0.368		0.951		0.426		0.309		0.291
Mean age (p.value)	16	-0.06 (0.790)	9	**0.46 (0.000)**	18	0.01 (0.939)	20	0.01 (0.939)	14	1.09 (0.603)
N intervention/month (p.value)	20	-0.02 (0.655)	12	-0.07 (0.681)	22	-0.02 (0.555)	24	-0.02 (0.555)	17	0.10 (0.849)
Males (%) (p.value)	18	0.03 (0.177)	12	-0.06 (0.467)	20	0.02 (0.058)	22	0.02 (0.058)	15	-0.17 (0.388)
Study size (p.value)	19	0.01 (0.362)	12	0.00 (0.700)	21	0.000 (0.408)	23	0.000 (0.408)	17	0.00 (0.888)

**Results in bold indicate statistical significance at 5% level.*

Figure S11. Funnel plot of BMI. (Eggers' test p.value=0.799)

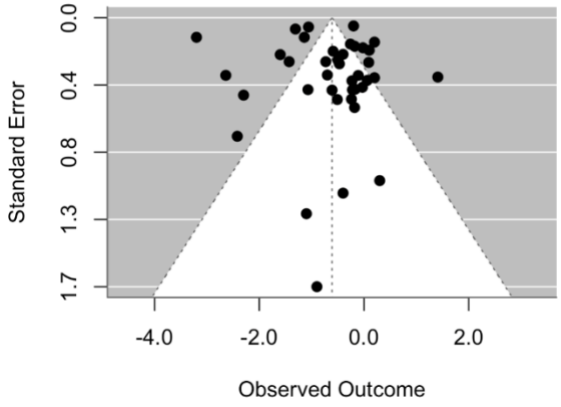

Figure S12. Funnel plot of weight. (Eggers' test p.value=0.845)

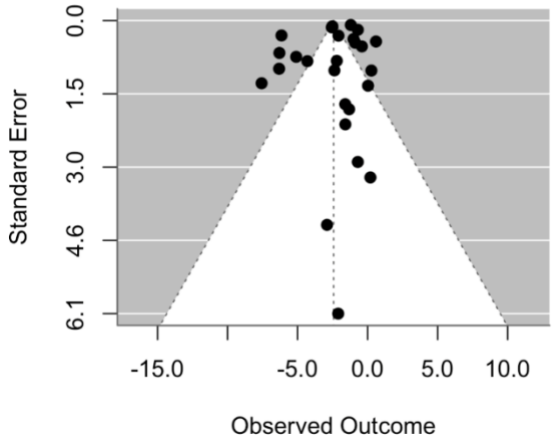

Figure S13. Funnel plot of body fat. (Eggers' test p.value=0.998)

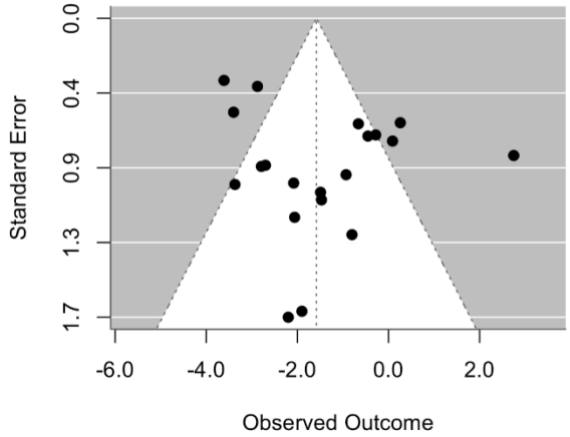

Figure S14. Funnel plot of total cholesterol. (Eggers' test p.value=0.711)

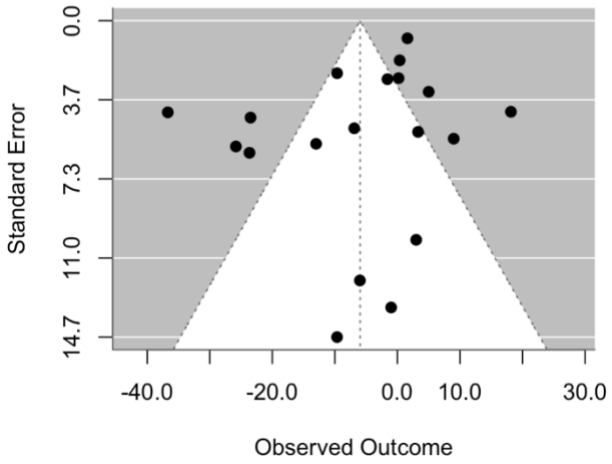

Figure S15. Funnel plot of HDL cholesterol. (Eggers' test p.value = 0.345)

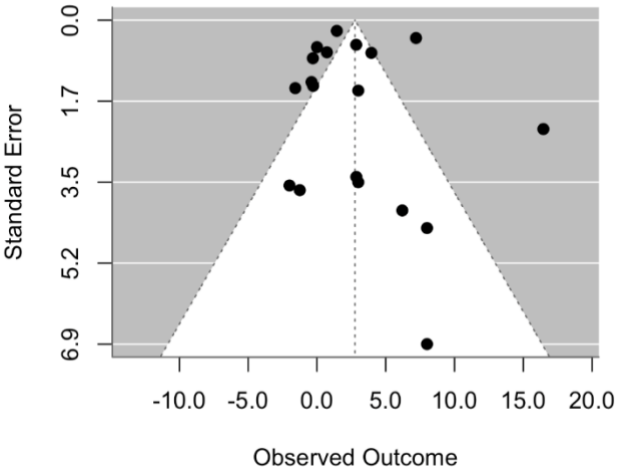

Figure S16. Funnel plot of LDL cholesterol. (Eggers' test p.value = 0.645)

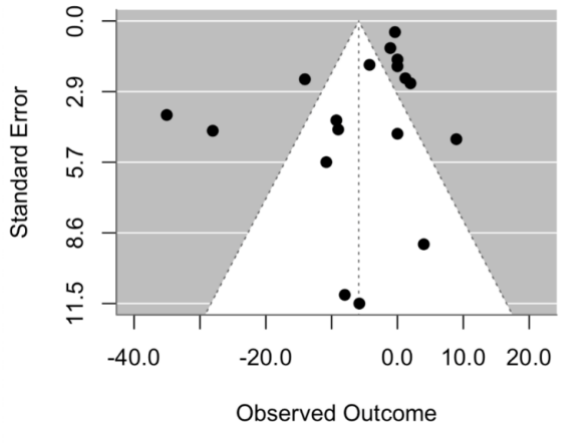

Figure S17. Funnel plot of DBP (Eggers' test p.value = 0.342)

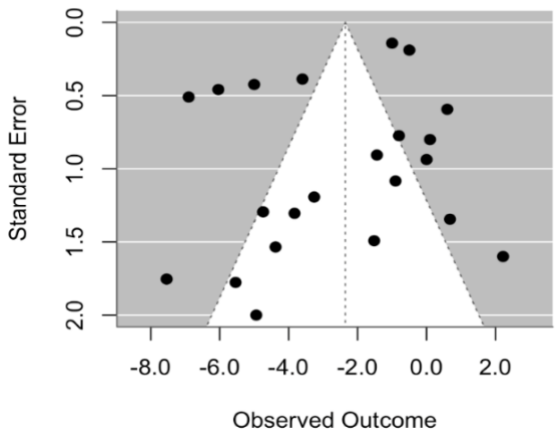

Figure S18. Funnel plot of waist circumference. (Eggers' test p.value = 0.330)

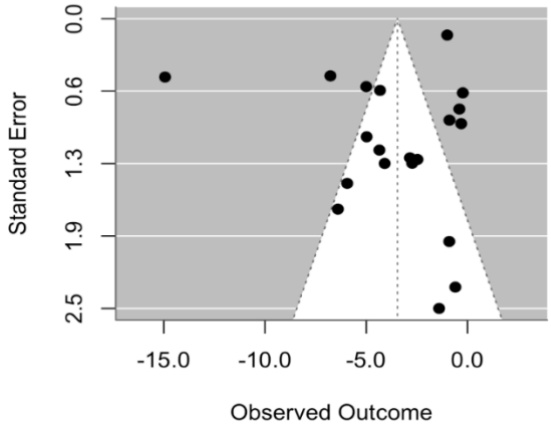

Figure S19. Funnel plot of SBP. (Eggers' test p.value = 0.447)

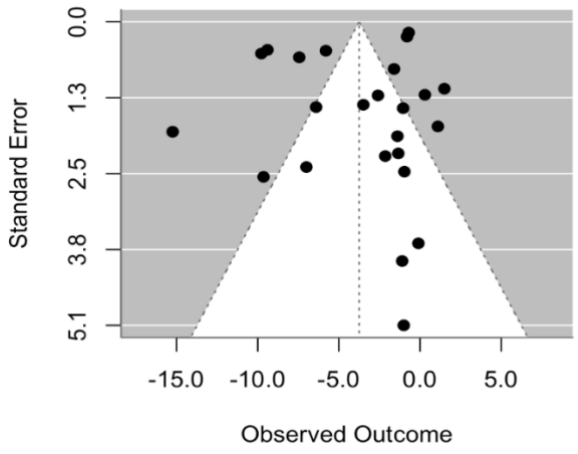

Figure S20. Funnel plot of triglycerides. (Eggers' test p.value=0.037)

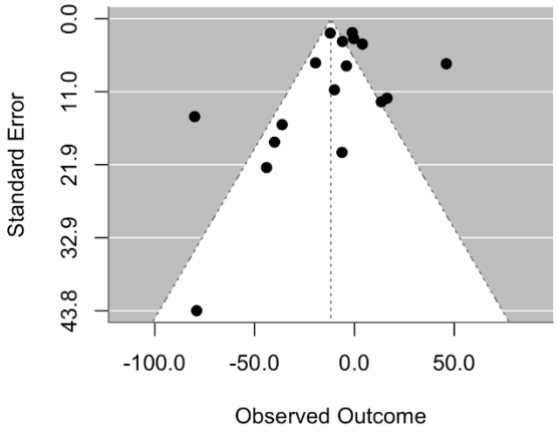

Figure S21. Funnel plot of glucose. (Eggers' test p.value=0.107)

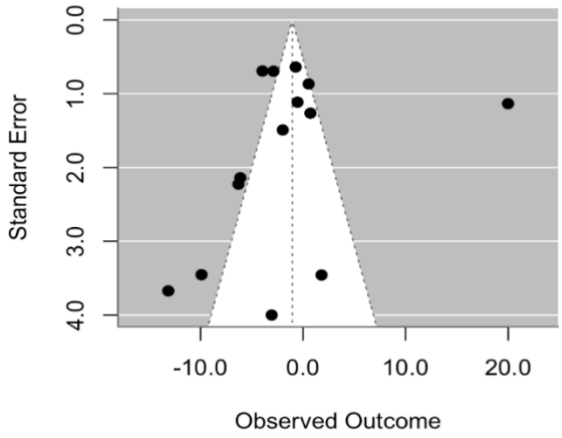

Figure S22. Funnel plot of smoking cessation. (Eggers' test p.value=0.006)

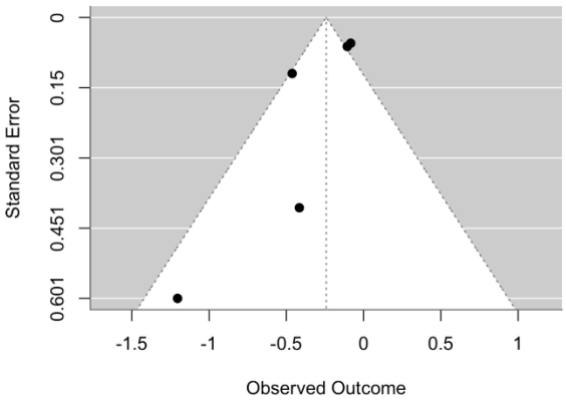

Table S9. Summary of GRADE ratings and justifications for downgrading.

	BMI: MODERATE ⊕⊕⊕◯		Weight: MODERATE ⊕⊕⊕◯		Body fat: HIGH ⊕⊕⊕⊕	
	Rating	**Rationale**	**Rating**	**Rationale**	**Rating**	**Rationale**
Risk of Bias	0	Most of the included studies (23/30) are of low risk of bias, and RCTs (20/30). Given the overall high methodological quality, we did not downgrade the evidence for risk of bias.	0	4/19 studies were at high risk of bias, particularly due to risk of confounding bias. However, most information is from studies at low risk of bias.and plausible bias is unlikely to seriously alter the results. No serious limitations do not downgrade.	0	3/16 studies were at high risk of bias, particularly due to risk of confounding bias. However, most information is from studies at low risk of bias.and plausible bias is unlikely to seriously alter the results. Hence, we did not downgrade the quality of evidence.
Indirectness	-1	Some studies (7/30) targeted individuals with specific conditions (obesity/hypertension/hypercholesterolemia). The focus on particular subgroups may reduce the applicability of the findings to a broader population. As such, the results may not fully represent individuals without these conditions, potentially limiting the generalizability and external validity of the evidence. Moreover, substantial heterogeneity was observed in subgroup analysis according to health status of participants. Thus, we downgraded for indirectedness.	-1	6 out of 19 studies targeted individuals with specific conditions (obesity/hypertension/hypercholesterolemia). The focus on particular subgroups may limit the generalizability and external validity of the evidence. A borderline statistical significance was found in subgroup analysis according to health status of participants. Thus, we downgraded for indirectedness.	0	4 out of 19 studies targeted individuals with specific conditions (obesity/hypertension/hypercholesterolemia). However, no statistically significant hererogeneity was found in in subgroup analysis according to health status of participants. Thus, we did not downgrade for indirectedness.
Inconsistency	0	We did not downgrade for inconsistency because the point estimates did not vary widely across studies, and there was overlap in the confidence intervals. Although the statistical test for heterogeneity was significant, the I^2 value of 31% indicates only moderate between-study heterogeneity, which is not large enough to warrant a downgrade for inconsistency.	0	Level 3 I^2 value is 20.7%, indicating low heterogeneity and although the point estimates varied across studies, the confidence intervals were overlapping. Hence, we did not downgrade for inconsistency.	0	Although between-study heterogeneity was high (63%), the point estimates remained relatively consistent across studies, and the confidence intervals showed substantial overlap. This suggests that despite statistical heterogeneity, the overall direction and magnitude of the effect were not highly variable, reducing concerns about inconsistency.
Imprecision	0	We did not downgrade for imprecision because the confidence intervals around the effect estimate were narrow.	0	We did not downgrade for imprecision because the confidence intervals around the effect estimate were not overly wide.	0	We did not downgrade for imprecision because the confidence intervals around the effect estimate were narrow.
Publication bias	0	The rating was not downgraded due to publication bias because the visual inspection of the funnel did not indicate concerning evidence of asymmetry, and the p.value of Egger test was not significant (p=0.799).	0	The rating was not downgraded due to publication bias because the visual inspection of the funnel plot did not indicate concerning evidence of asymmetry, and the p.value of Egger test was not significant (p=0.845).	0	The rating was not downgraded due to publication bias because the visual inspection of the funnel plot did not indicate concerning evidence of asymmetry, and the p.value of Egger test was not significant (p=0.998).
Large magnitude	0	While the results suggest a notable effect, the possible presence of indirectedness prevents us from rating up the quality of evidence.	0	While the effect size appears meaningful, indirectedness prevents an upgrade for large magnitude of effect.	0	The certainty of the evidence is already rated as high, thus no further upgrading is required. Moreover, it is difficult to unambiguously determine what constitutes a large magnitude of effect in terms of body fat reduction.
Dose response	0	The included studies do not systematically examine different levels of intervention intensity. Even if some studies provided an intervention-response relationship, the heterogeneity across studies makes it difficult to detect a consistent pattern.	0	The included studies do not systematically assess different levels of intervention intensity to establish a dose-response pattern.	0	The included studies do not systematically assess different levels of intervention intensity to establish a dose-response pattern.
Residual confounding	0	We did not upgrade the evidence for plausible residual confounding because there is no indication that residual confounding is likely to have substantially reduced effect.	0	There is no strong indication that residual confounding would have led to an underestimation of the effect.	0	There is no strong indication that residual confounding would have led to an underestimation of the effect.

	Waist Circumference: MODERATE ⊕⊕⊕◯		Total Cholesterol: VERY LOW ⊕◯◯◯		HDL cholesterol MODERATE ⊕⊕⊕◯	
	Rating	**Rationale**	**Rating**	**Rationale**	**Rating**	**Rationale**
Risk of Bias	0	6/17 studies were at high risk of bias. However, most information is from studies at low risk of bias.and plausible bias is unlikely to seriously alter the results. No serious limitations do not downgrade.	0	4/16 studies were at high risk of bias, especially due to a lack of control for confounding. However most evidence came from studies at low risk of bias studies and plausible bias is unlikely to alter results.	0	5/16 studies were at high risk of bias, especially due to a lack of control for confounding. However most evidence came from studies at low risk of bias studies and plausible bias is unlikely to alter results.
Indirectness	-1	Four studies targeted individuals with specific conditions (obesity/hypertension/hypercolesterolemia). Subgroup analysis according to health conditions of participants revealed signifcant heterogeneity. As a result, we downgraded the quality of evidence for indirectness.	-1	Only two studies focused solely on obese individuals. However, significant difference in subgroup analysis supports the downgrade for indirectedness.	0	Three studies focused solely on obese individuals. No significant difference in subgroup analysis was found. Hence, we did not downgrade for indirectedness.
Inconsistency	0	We did not downgrade for inconsistency because the point estimates did not vary widely across studies, and there was overlap in the confidence intervals. Although the statistical test for heterogeneity was significant, the I^2 value of 36% indicates only moderate between-study heterogeneity, which is not large enough to warrant a downgrade for inconsistency.	-1	The point estimates varied greatly across studies, suggesting inconsistency in the effect. Hence, we downgraded by one level	-1	We downgraded due to inconsistency because of high between study heterogeneity (I^2 =96%) and consistent variability across point estimates.
Imprecision	0	We did not downgrade for imprecision because the confidence interval around the effect estimate was narrow.	-1	We downgraded the quality of evidence for imprecision because the confidence intervals around the effect estimate were wide, indicating uncertainty in the precision of the results.	0	We did not downgrade for imprecision because the confidence interval around the effect estimate was narrow.
Publication bias	0	The rating was not downgraded for publication bias because visual inspection of the funnel plot did not show significant asymmetry, except for one outlier. Additionally, the Egger test did not yield a significant p-value (p = 0.330), further supporting the absence of publication bias.	0	The rating was not downgraded due to publication bias because the visual inspection of the funnel did not indicate concerning evidence of asymmetry, and the p.value of Egger test was not significant (p=0.711).	0	The rating was not downgraded due to publication bias because the visual inspection of the funnel did not indicate concerning evidence of asymmetry, and the p.value of Egger test was not significant (p=0.345).
Large magnitude	0	While the results suggest a notable effect, the possible presence of risk of bias and indirectedness prevents us from rating up the quality of evidence.	0	There weren't sufficient conditions to support a rating up of the quality of evidence.	0	There weren't sufficient conditions to support a rating up of the quality of evidence.
Dose response	0	The included studies do not systematically examine different levels of intervention intensity.	0	The included studies do not systematically examine different levels of intervention intensity. Even if some studies provided an intervention-response relationship, the heterogeneity across studies makes it difficult to detect a consistent pattern.	0	The included studies do not systematically examine different levels of intervention intensity.
Residual confounding	0	We did not upgrade the evidence for plausible residual confounding because there is no indication that residual confounding is likely to have substantially reduced effect.	0	We did not upgrade the evidence for plausible residual confounding because there is no indication that residual confounding is likely to have substantially reduced effect.	0	We did not upgrade the evidence for plausible residual confounding because there is no indication that residual confounding is likely to have substantially reduced effect.

	LDL cholesterol: VERY LOW ⊕◯◯◯		DBP: HIGH ⊕⊕⊕⊕		SBP: HIGH⊕⊕⊕⊕	
	Rating	**Rationale**	**Rating**	**Rationale**	**Rating**	**Rationale**
Risk of Bias	0	5/15 studies were at high risk of bias, especially due to lack of control for confounding. However, since most studies were at low risk of bias we did not downgrade the quality of evidence.	0	5/19 studies were at high risk of bias, particularly due to risk of confounding bias. However, we did not downgrade the quality of evidence because most studies were at low risk of bias and plausible bias would unlikely affect the results.	0	5/20 studies were at high risk of bias, particularly due to risk of confounding bias. However, we did not downgrade the quality of evidence because most studies were at low risk of bias and plausible bias would unlikely affect the results.
Indirectness	-1	3/15 studies focused solely on obese individuals. We also found significant differences in subgroup analysis by the health status of participants. Hence, we did not downgrade for indirectedness.	0	Four studies targeted individuals with specific conditions (obesity/hypertension/hypercholesterolemia). Subgroup analysis according to health conditions of participants did not reveal significant heterogeneity. As a result, we did not downgrade the quality of evidence for indirectness	0	Four studies targeted individuals with specific conditions (obesity/hypertension/hypercholesterolemia). Subgroup analysis according to health conditions of participants did not reveal significant heterogeneity. As a result, we did not downgrade the quality of evidence for indirectness
Inconsistency	-1	Between-study heterogeneity was low (I^2 = 25%). However, the consistent variability across point estimates indicated potential inconsistency in the direction and magnitude of effects. As a result, the certainty of evidence was downgraded for inconsistency.	0	Between study heterogeneity was high ((I^2 =94%). However, point estimates were consisent and confidence intervals were overlapping. Hence no dowgrade was made.	0	Between study heterogeneity was high (I^2 =84%). However, point estimates were consisent and confidence intervals were overlapping. Hence no dowgrade was made.
Imprecision	-1	We downgraded the quality of evidence for imprecision because the confidence intervals around the effect estimate were wide, indicating uncertainty in the precision of the results.	0	We did not downgrade for imprecision because the confidence intervals around the effect estimate were narrow.	0	We did not downgrade for imprecision because the confidence intervals around the effect estimate were narrow.
Publication bias	0	The rating was not downgraded due to publication bias since the visual inspection of the funnel did not indicate concerning evidence of asymmetry, and the p.value of Egger test was not significant (p=0.645).	0	The rating was not downgraded due to publication bias because the visual inspection of the funnel did not indicate concerning evidence of asymmetry, and the p.value of Egger test was not significant (p=0.342).	0	The rating was not downgraded due to publication bias because the visual inspection of the funnel did not indicate concerning evidence of asymmetry, and the p.value of Egger test was not significant (p=0.447).
Large magnitude	0	There weren't sufficient conditions to support a rating up of the quality of evidence.	0	The certainty of the evidence is already rated as high, thus no further upgrading is required. Moreover, it is difficult to unambiguously determine what constitutes a large magnitude of effect in terms of blood pressure.	0	The certainty of the evidence is already rated as high, thus no further upgrading is required. Moreover, it is difficult to unambiguously determine what constitutes a large magnitude of effect in terms of blood pressure.
Dose response	0	The included studies do not systematically examine different levels of intervention intensity.	0	The included studies do not systematically examine different levels of intervention intensity.	0	The included studies do not systematically examine different levels of intervention intensity.
Residual confounding	0	We did not upgrade the evidence for plausible residual confounding because there is no indication that residual confounding is likely to have substantially reduced effect.	0	We did not upgrade the evidence for plausible residual confounding because there is no indication that residual confounding is likely to have substantially reduced effect.	0	We did not upgrade the evidence for plausible residual confounding because there is no indication that residual confounding is likely to have substantially reduced effect.

	Glucose: LOW ⊕⊕◯◯		Tryglycerides: VERY LOW ⊕◯◯◯		Smoking: LOW ⊕⊕◯◯	
	Rating	**Rationale**	**Rating**	**Rationale**	**Rating**	**Rationale**
Risk of Bias	0	3/11 studies were rated as having a high risk of bias. Most evidence came from studies with low risk of bias. Consequently, evidence was not downgraded for risk of bias.	0	4/11 studies were rated as having a high risk of bias. Most evidence came from studies with low risk of bias. Consequently, evidence was not downgraded for risk of bias.	0	Only one out of five studies was assessed as having a high risk of bias. Therefore, the certainty of evidence was not downgraded for risk of bias.
Indirectness	0	3/11 studies targeted individuals with specific health conditions (e.g., obesity, hypertension, or hypercholesterolemia). However, subgroup analyses based on participants' health status did not reveal significant heterogeneity in effect estimates. Therefore, the certainty of evidence was not downgraded for indirectness	0	3/11 studies targeted individuals with specific health conditions (e.g., obesity, hypertension, or hypercholesterolemia). Subgroup analyses on participants' health status did not reveal significant heterogeneity in effect estimates. Therefore, the certainty of evidence was not downgraded for indirectness	-1	A downgrade was applied due to indirectness, as there were substantial differences in the nature of the smoking interventions.
Inconsistency	-1	Between-study heterogeneity was high (I^2 = 97%). Moreover, there was consistent variability across point estimates, and confidence intervals were not consistently overlapping. As a result, the certainty of evidence was downgraded for inconsistency.	-1	Between-study heterogeneity was high (I^2 = 97%), and point estimates varied greatly, with confidence intervals not overlapping. As a result, the certainty of evidence was downgraded for inconsistency.	0	Between-study heterogeneity was high (I^2 = 77%), indicating substantial variability in the results. Despite this, all point estimates across the studies suggested a reduction in smoking, and the confidence intervals overlapped, indicating a consistent trend in the overall results. Therefore, the certainty of the evidence was not downgraded for inconsisteny.
Imprecision	-1	We downgraded the quality of evidence for imprecision because the confidence intervals around the effect estimate were wide, indicating uncertainty in the precision of the results.	-1	We downgraded for imprecision because the confidence intervals around the effect estimate were wide.	0	We did not downgrade for imprecision because the confidence intervals around the effect estimate were narrow.
Publication bias	0	The rating was not downgraded due to publication bias. Although some asymmetry was observed in the funnel plot, it was not considered substantial. Moreover, the Egger test did not indicate significant small study effects (p = 0.107), supporting the decision not to downgrade the certainty of evidence for publication bias.	-1	The rating was downgraded due to publication bias because the visual inspection of the funnel indicated evidence of asymmetry, and the p.value of Egger test was significant (p=0.037).	-1	The rating was downgraded due to publication bias because the visual inspection of the funnel indicated evidence of asymmetry, and the p.value of Egger test was significant (p=0.006).
Large magnitude	0	The pooled effect size was not significant.	0	The pooled effect size was not significant.	0	There weren't sufficient conditions to support a rating up of the quality of evidence.
Dose response	0	The included studies do not systematically examine different levels of intervention intensity.	0	The included studies do not systematically examine different levels of intervention intensity.	0	The included studies do not systematically examine different levels of intervention intensity.
Residual confounding	0	We did not upgrade the evidence for plausible residual confounding because there is no indication that residual confounding is likely to have substantially reduced effect.	0	We did not upgrade the evidence for plausible residual confounding because there is no indication that residual confounding is likely to have substantially reduced effect.	0	We did not upgrade the evidence for plausible residual confounding because there is no indication that residual confounding is likely to have substantially reduced effect.

Table S10. PRISMA Checklist.

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	p. 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	p. 1
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	p. 3
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	p. 3
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	p. 4
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	p. 3
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	p. 1 (Supplementary material)
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	p. 4
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	p. 4
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	p. 5
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	p. 4-5
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	p. 5
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	p. 6
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	p. 5-6
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data	p. 5-6

		conversions.	
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	p. 5-6
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	p. 5-6
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	p. 5-6
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	p. 5-6
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	p. 5-6
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	p. 5
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	p. 6
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	p. 6
Study characteristics	17	Cite each included study and present its characteristics.	p. 6-10
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	p. 9-11
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	p. 12-13
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	p. 12-13
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	p. 12-13
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	p. 14
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	p. 12-13
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	N/A
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	p. 1
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	p. 15
	23b	Discuss any limitations of the evidence included in the review.	p. 17

	23c	Discuss any limitations of the review processes used.	p. N/A
	23d	Discuss implications of the results for practice, policy, and future research.	p. 17-18
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	p. 3
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	N/A
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	N/A
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	p.
Competing interests	26	Declare any competing interests of review authors.	p. 18
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	p. 18

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021; 372:n71. doi: 10.1136/bmj.n71.

La Medicina del Lavoro

Original article

Med. Lav. 2025; 116 (6): 17437
DOI: 10.23749/mdl.v116i6.17437

Evaluation of the Hearing Function in the Orchestra Professional Musicians of the Teatro alla Scala in Milan

Terenzio Cassina[1,*], Alessia Tomasin[2], Orietta Calcinoni[1], Umberto Ambrosetti[2], Eliana Filipponi[3], Valeria Castelli[4], Riccardo Bianchin[2], Valentina Teodorani[2], Laura Magnoni[4], Leonardo Braga[5], Linda Parati[6], Federica Di Berardino[2,4]

[1] Fondazione Teatro alla Scala di Milano
[2] Department of Clinical Sciences and Community Health, Dipartimento di Eccellenza 2023–2027, University of Milan, Milan, Italy
[3] Direzione Aziendale Professioni Sanitarie, Fondazione IRCCS Ca' Grande Ospedale Maggiore Policlinico
[4] Audiology Unit, Department of Surgical Sciences, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico
[5] Laboratory of Healthcare Research & Pharmacoepidemiology, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy
[6] Acoustical Consultant, Acustica Parati & Co, Crema, Italy

Keywords: Opera; Music; Orchestral Professional Musicians; Musicians Hearing Threshold; Hypoacusis; Hearing Disorders; Occupational Workers; Noise-Induced Hearing Loss

Abstract

Background: *This study was conducted to determine whether occupational exposure to high sound levels, typical of an opera orchestra, can cause hearing loss.* **Methods:** *The orchestra professors at Teatro alla Scala in Milan underwent ear examinations, pure-tone audiometry, and other audiological tests. The hearing thresholds of these musicians were compared with those of populations not exposed to occupational noise and with populations exposed to industrial noise. Noise exposure levels were estimated through a phonometric survey conducted at our theater in 2011, which largely confirmed the exposure levels outlined in European guidelines.* **Results:** *The average audiometric thresholds of the orchestra musicians were slightly worse than the median thresholds of a healthy, non-noise-exposed population. In three subjects (2.8%), bilateral hearing impairment (PTA 0.5, 1, 2, 3, 4 kHz > 25 dB HL) was observed; in four violinists (3.7%) left unilateral hearing impairment was found. This rate is lower than the expected risk from similar industrial noise exposures. Comparing these audiometries with those from about ten years earlier shows that the hearing threshold decline in the study group is comparable to that caused by presbycusis.* **Conclusions:** *The risk of noise-induced hearing loss among professional orchestra musicians appears lower than predicted by the UNI ISO 1999:2015 standard. A few cases of hearing loss due to chronic acoustic trauma were noted, particularly among violinists who demonstrated a higher incidence of left unilateral hearing loss. The high levels of sound exposure and the presence of some hearing loss cases highlight the need for targeted preventative measures in this work activity.*

Received 28.08.2025 – Accepted 22.11.2025
*Corresponding Author: Terenzio Cassina; E-mail: terenziocassina@virgilio.it

1. Introduction

Prolonged exposure to high sound levels can lead to noise-induced hearing loss (NIHL). NIHL depends on the level and duration of exposure but also factors like age, medication use, pre-existing or concurrent ear diseases [1, 2], individual susceptibility, and genetics. Zhang et al. [3] showed that genetic susceptibility significantly influences NIHL occurrence.

A literature search on PubMed focusing on noise-induced damage in professional classical orchestra musicians identified 67 articles published from 1960 to 2025. Regarding risk, the literature generally agrees that professors of opera orchestras are exposed to high sound levels. Exposure levels are often above 85 dBA, and the use of hearing protectors (PPE) is uncommon in this artistic environment [4]. The European Guidelines for preventing hearing damage in the performing arts sector, implemented in Italy [5-6], indicate that weekly exposure exceeds 85 dBA for all instruments except the double bass; for some wind and percussion instruments the intensity is around 90 dBA. Concerning potential NIHL among music performers [7], especially among professional orchestra musicians, the existing literature is inconsistent, with differences in study populations and exposure durations. Hake [4] evaluated 200 orchestral players, of whom only 58 were involved in classical opera or symphony orchestras; the rest were students from academies, conservatories, or amateurs, with limited details on age, exposure time, and extent. Moore et al. [8] and Quian et al. [9] proposed corrective assessment criteria, particularly regarding exposure times. Exposure is rarely documented through work schedules and repertoire. Several studies have focused solely on exposure during live performances [10-11]. Others examine exposure during individual practice, which often occurs in suboptimal acoustic environments, such as standard classrooms [12]. In the literature, the concept of 'damage' is defined differently: for some authors it is the loss of 20 dB HL at a single frequency [13]. Conversely, others define it as an assessment relative to the expected loss per population. Some studies assess NIHL not through audiometric tests but using specially designed questionnaires [14], or through surveys generated by dedicated apps [4].

In the review by Behar et al. [15], challenges were identified in drawing general conclusions about the occupational risks faced by musicians due to inconsistent methods across studies, both in instrumentation and procedures, as well as variability in reported performance times. To assess the impact of noise on hearing, proper tests must be conducted following a sufficient rest period since the last exposure to ensure the complete reversibility of the temporary hearing threshold shift. Good clinical practice recommends an acoustic rest period of at least 16 hours before audiometric testing [16]. However, in musicians' hearing evaluations, maintaining this optimal acoustic rest time is often challenging because of logistical issues such as work schedules and the habit of continuous practice and preparation. Nonetheless, Behar [17] suggests that classical music can cause a threshold shift of approximately 11 dBHL, which recovers in about 55 minutes. Due to these variables, research on noise-induced hearing loss (NIHL) among orchestral musicians has produced inconsistent findings. Most studies do not report an increased risk of hearing loss, even though these musicians are exposed to high noise levels but typically do not wear hearing protectors during performances, unlike industrial workers. Recent research has also pointed to the possible presence of hearing disorders such as tinnitus, hyperacusis, and sound perception distortions (e.g., diplacusis), which may impair a musician's artistic abilities. Hyperacusis is defined as "abnormal sensitivity to everyday sounds or noises," which can make certain sounds seem "painfully loud" [18-19], regardless of hearing threshold [20-21], and can cause difficulty perceiving loudness, a situation that can be especially disorienting for an orchestra conductor or professor. Diplacusis involves pitch perception difficulties or distortions, which can severely impact a musician's career—particularly for violinists or conductors. Sensorineural deafness at high frequencies, often linked with tinnitus, may lead a musician to play too loudly, with disastrous effects on performance quality [22-23]. This study aims to make a significant contribution to evaluating the risk of hearing loss among classical orchestra musicians, based on

documented exposure levels, and to identify prevention strategies tailored for this specific context.

2. Methods

2.1. Fieldwork

The study involved 119 professional musicians from the orchestra of the Fondazione Teatro alla Scala in Milan. Each participant was informed about the study's purpose and procedures and signed an informed consent form. They also received a document regarding personal data processing (all forms are available on the Teatro alla Scala Foundation website). The Ethics Committee of the University of Milan gave a positive opinion on this study and approved the informed consent form during its meeting n. 48.23 on April 18, 2023. The study's methods and purpose were also explained to the management of the Fondazione Teatro alla Scala and the union representatives of the orchestral musicians. All procedures in this study adhere to relevant international and institutional ethical standards on human research and follow the Helsinki Declaration of 1975, revised in 2008. Participation was voluntary, and written informed consent to participate and publish the results was obtained from all participants. Their anonymity has been preserved.

2.2. Noise Exposure Assessment

The phonometric survey conducted at our theatre in 2011 largely confirmed the exposure data outlined in the European guidelines. The evaluation process adopted accounted for all aspects that influence sound exposure. After consulting with an artistic commission, three works were selected based on their acoustic load on the musicians: The Magic Flute (Light), Tosca (Medium), and Turandot (Heavy). The musicians were grouped by instrument family, except for the strings, due to the large number of instruments and their specific positions in the orchestra. During the measurements, carried out both in the pit during performances and rehearsals, and in the rehearsal room, sound level meters and dosimeters positioned on the same musicians were used. For The Magic Flute, two dosimeters and seven sound level meters were used, recording a total of 29 positions; for Tosca, three dosimeters and six sound level meters, with 27 recorded positions; and for Turandot, three dosimeters and eight sound level meters, totaling 34 recorded positions. Individual practice is a fundamental part of a musician's activity, therefore separate measurements were also performed on some instruments such as the violin, cello, double bass, oboe, and clarinet. The activity duration for each was determined by cross-referencing data from the theatre management and the musicians. It was found that 33.3% of working time was spent in the rehearsal room, 42.4% in the orchestra pit, and 24.2% in individual practice. Furthermore, musicians engage with different repertoires within the same week, performing various operas while preparing a new one. For this reason, for each instrument or family, different weekly noise exposure levels (Lex,w) were established, combining different operas. For example, for the violin, the minimum Lex,w involves 2 light works and 1 heavy work, while the maximum involves 2 medium operas and 1 heavy work. Estimated weekly exposure values during orchestra activities are summarized in Table 1. In conclusion, musicians are exposed to Lex,w during different working weeks, with variations across instruments or families within 2-3 dB(A). The measurements from the individual practice showed a marked difference in exposure for the violin between the left ear (92.2 dB(A)) and the right ear (81.4 dB(A)), while the difference for other instruments was much smaller or nonexistent.

2.3. Audiological Assessment

In the auditory evaluation protocol, each participant (as part of the health surveillance activities arranged by the Competent Doctor in accordance with DL n.81/2008 [24]) underwent a specialist audiological assessment lasting approximately 30 minutes, which included:

- anamnesis and subjective evaluation of one's hearing condition;
- specialistic audiological examination to detect tympanic objectivity (video-otoscopy performed by Interacoustics Video Otoscope mod. VIOT);

Table 1. Results of the phonometric survey carried out by the Teatro alla Scala Foundation in Milan. The LEX,w is estimated for three levels of acoustic load perceived by musicians for the type of artistic production performed, and the dosimetric data [Leq dB(A)] measured for single instruments during an individual study with the dosimeter microphone placed on both sides of the neck are indicated.

Instruments	Estimated weekly exposure values *[$L_{EX,w}$ dB(A)]* during orchestra activities			Exposure levels *[L_{eq} dB(A)]* during individual study measured with dosimeters worn on both sides of the neck	
	Mild	Medium	High	Left side	Right side
Violin	85.5	85.8	87.3	92.2	81.4
Viola	86.9	88.0	89.8	NV	NV
Cello	85.6	86.8	87.1	87.0	87.6
Double bass	79.2	80.1	80.8	81.3	81.4
Woodwinds	87.7	87.9	89.5	87.7 *(Oboe)*	86.9
				80.4 *(Clarinet)*	77.7
Brass	90.2	91.2	92.2	92.3 *(Horn)*	93.6
				90.7 *(Trumpet)*	93.4
Percussions	92.5	92.9	93.3	NV	NV

- tympanometry, to assess the function of the tympanic-ossicular system (Interacoustics Impedance Meter mod. AT235), the stapedial reflex was not recorded to avoid the onset of tinnitus, a rare occurrence but always possible given the high intensity of the acoustic stimulus according to the ANSI S3.39-1987;
- liminal threshold pure-tone audiometry (Interacoustics Clinical Audiometer mod. AC 40B) in a silent soundproof booth (UNI EN ISO 8253 1:2010) according to the ANSI S3.1-1999; ANSI S3.6-2018;
- study of acoustic otoemissions from distortion products (DPOAE), (Interacoustics Otoemission System mod. Lyra DP+TE);
- Dichotic Digit Test (DDT), to assess the integration of auditory information binaurally.

Participants were also asked to complete an online questionnaire based on Laitinen's [14], which covered various topics related to the musicians' working conditions, auditory issues (such as tinnitus, hyperacusis, sound distortions, diplacusis, and the sensation of occlusion), use of hearing protection devices, and difficulties experienced with their use. Audiometric tracings were assessed and classified according to the Merluzzi-Pira-Bosio method, as outlined in the SIMLII's "Guidelines for the prevention of hearing damage from noise in the workplace" [25]. The longitudinal assessment of the musicians' hearing thresholds involved comparing current audiometric tracings with those performed by the health service of Teatro alla Scala for the health surveillance of orchestra professional musicians from 2011 to 2019.

2.4. Statistical Analysis

Data were reported in a Microsoft Excel (v16.98) datasheet for dataset management and graph processing. The analysis involved correlating threshold pure tone audiometric values with personal variables, sound exposure from musical instrument use, and occupational seniority in the orchestra. The Normality Test (Shapiro-Wilk)[27], which returns a p-value > 0.05 if the data follow a normal distribution and a p-value < 0.05 if the data are non-normal, was used. To determine whether differences between groups were statistically significant, the Wilcoxon test was applied if the normality assumption was not met. If the data were normally distributed, the t-test was used instead. The statistical analyses were performed using RStudio (version 4.3.1). The comparison with the UNI ISO 1999:2015 standard [26] was conducted through the use of the software "Rumours.21" by

Casini S., which was used to assess the risk of hearing damage using the HTLAN (Hearing Threshold Level associated with Age and Noise) protocol.

3. Results

The study was conducted on 107 musicians because, out of the total population of 119 professors at the orchestra of the Fondazione Teatro alla Scala in Milan, one refused to participate, and 11 individuals were excluded due to unreliable and non-reproducible audiometric tests (3 individuals) or the presence of damage unrelated to noise exposure (8 individuals).

The main characteristics of the study subjects are shown in Table 2: 83 were males (78%) and 24 females (22%), with a mean age of 48.0 ± 10 years, and an average length of service in the orchestra of 27.1 ± 10.7 years. The average weekly noise exposure, based on different instruments and categorized according to the risk bands specified in Article 189 of DL n. 81/2008, is also illustrated.

3.1. Comparison With the Population Not Occupationally Exposed to Noise

Audiometric tests were collected from 6 groups, divided by gender and age. For each group, the average audiometric result was compared with the hearing threshold tracing for the group's average age, as reported in the ISO standards. Comparisons are shown in Figures 1a and 1b for male and female subjects, respectively.

1. with the standard of ISO 7029 taken from Database A of ISO 1999/2015, which represents "the statistical distribution of the hearing threshold of selected otologically normal populations (without signs or symptoms of hearing pathology) free of obstructive ear wax and without a history of occupational noise exposure". The threshold values are referred to as HTLA (Hearing Threshold Levels associated with Age)
2. with the statistical distribution provided by the B3 Database of the ISO 1999/2015 standard representing data from "an unselected population from an industrialized nation (Norway), otologically healthy and free of occupational noise exposure".

In all groups, the average thresholds of the orchestra's professional musicians at Teatro alla Scala were slightly worse than the 50th percentile of otologically healthy, non-noise-exposed populations, according to ISO 7029 and ISO 1999:2015 (Databases A and B3). The difference decreases with increasing age, while it is more pronounced in females up to age 50.

3.2. Assessment of the Incidence of Noise Damage

Audiometric tests were classified using the Merluzzi-Pira-Bosio (MPB) method, analyzing each ear separately to account for asymmetric exposure,

Table 2. Main characteristics of the sample and subdivision of the instruments by sound exposure class.

Subjects	Nr (%)	Age (years) Mean (range)	Weekly exposure $[L_{EX,w}$ dB(A)] Mean ± SD	Length of time (years) as orchestral musician (Mean ± SD)
Males	83 (78)	49 (23-62)	82.2 ± 3.6	27.7 ±10.8
Females	24 (22)	47 (29-62)	86.6 ± 1.3	24.9 ± 10.1
Instruments				
Double bass Keyboard	12(11.2)	47.6 (27-61)	< 85 dB	24.9 ± 9.7
Violin, Cello	46 (43)	49.5 (23-61)	85-87 dB	27.2 ± 10.4
Harp, Viola Woodwinds	32 (30)	51.7 (27-62)	87-90 dB	30.3 ± 8.6
Brass Percussions	16 (15)	43.6 (28-60)	>90 dB	22.2 ± 13.9

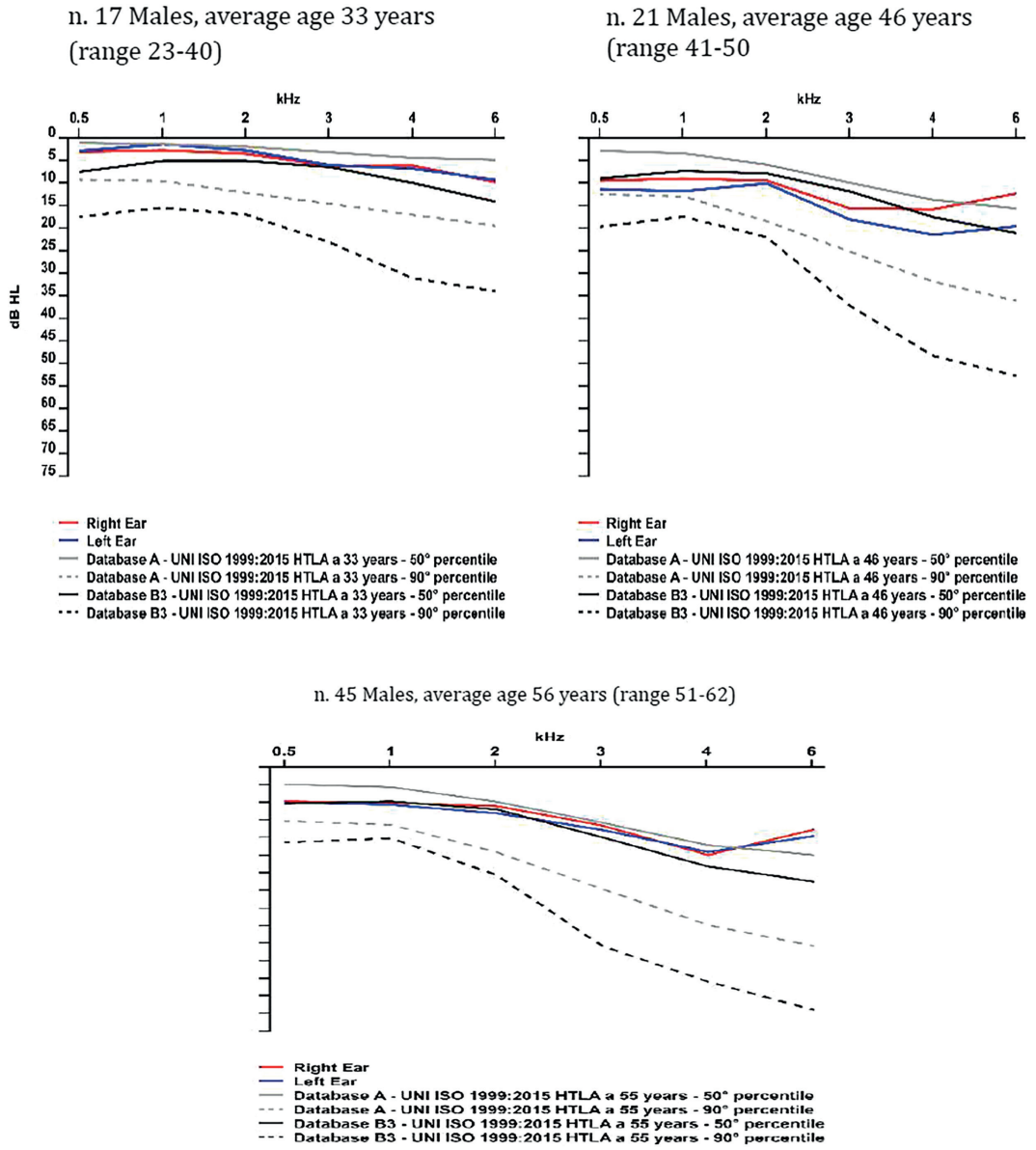

Figure 1a. Comparison of the average audiometric curves of the male population divided by age group with those of an otologically healthy reference population not exposed to noise (database A and B3 of the UNI ISO 1999/2015 standard.

especially among violinists, since asymmetric exposure to asymmetric instruments has been reported in the literature and was confirmed by the phonometric survey conducted at Teatro alla Scala.

Normal hearing (classes 0 and 1 A) was detected in 66.3 % of cases; the percentage of age-related hearing loss (class 8, presbycusis) was 10.3 %, while noise hearing impairment of varying degrees (classes 2 to 6, mono or bilateral) occurred in 23.3 % of the subjects. Noise hearing deficit was present in 25.3% of male musicians and 16.6% of female musicians;

- 10 subjects presented bilateral deficits, of which 50% were slightly asymmetrical,
- 8 have right-sided unilateral deficits, all of which are minor-class 2A,
- 7, of which 6 violinists, have left-sided hearing loss.

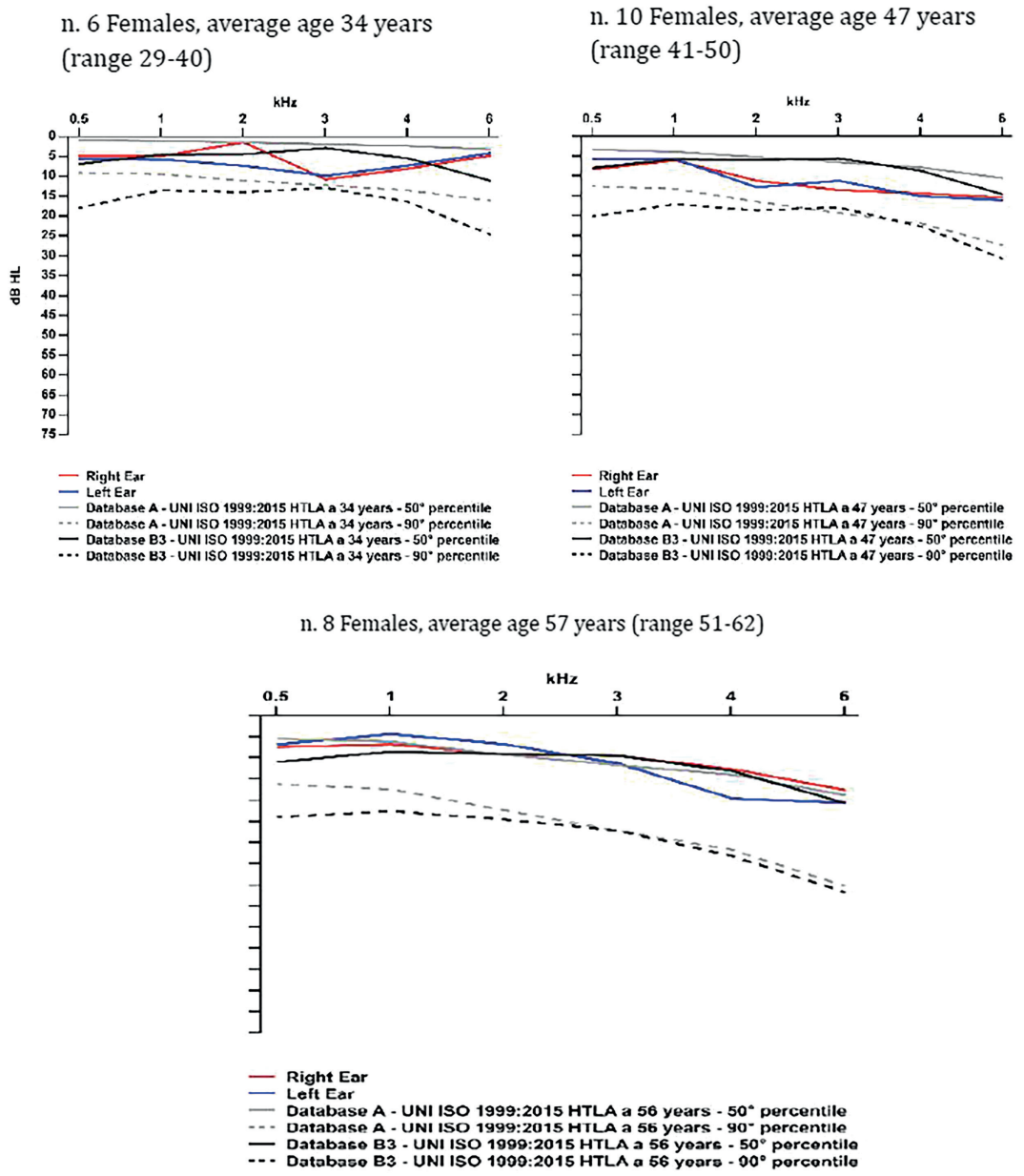

Figure 1b. Comparison of the average audiometric curves of the female population divided by age group with those of an otologically healthy reference population not exposed to noise (database A and B3 of the UNI ISO 1999/2015 standard.

According to the Merluzzi-Pira-Bosio (MPB) method, class 2 subjects have minimal hearing impairment, without impairment of comfortable social audibility, and usually without a subjective sensation of hearing impairment. In contrast, class 3A or higher includes hearing impairments due to chronic acoustic trauma, with increasing difficulty in understanding conversational speech.

About hearing impairments of class 3A or higher we found:

- 5 subjects with bilateral hearing loss, 2 violas, 1 flute, 1 percussion (with Lex-w > 87 dBA), and one violinist (with Lex-w between 85 and 87 dBA);
- 6 left unilateral hearing impairments, all in violinists.

These results confirm the expectation that violinists experience greater NIHL on the left side. There is a statistically significant difference in hearing thresholds between the left and right ears of violinists, with the left ear often showing poorer hearing only for these musicians. In fact, the comparison between the left/right hearing thresholds of violinists and those of musicians playing symmetrical instruments was also statistically significant (p=0.02) (Figure 2). This aligns with the phonometric evaluation conducted (Table 1): for violinists, during individual practice, the sound in the left ear is more than 10 dBA louder than in the contralateral ear.

3.3. Comparison with ISO 1999:2015 Noise-Induced Hearing Loss (NIHL) Risk

The risk of hearing damage in our sample was compared with that expected for workers exposed to industrial noise of the same intensity and duration, using ISO 1999:2015 as a reference. This standard with the HTLAN (Hearing Threshold Level associated with Age and Noise) values, indicates hearing threshold levels related to age and industrial noise exposure (intensity and years of exposure). We used the average biological damage indicator, dB 25HL, as a reference value, considering hearing damage present when the average threshold elevation at the 5 frequencies (0.5, 1, 2, 3, 4 KHz) is equal to or greater than 25 dB HL. We then calculated the risk as the percentage of the population with a hearing threshold at or above this value.

For a group of workers exposed to industrial noise of the same intensity and duration as our musicians, the expected risk of damage (PTA 0.5, 1, 2, 3, 4 KHz > 25 dB HTL) is 10%. In our sample, the actual cases of noise damage were lower, at 6.5%, with three subjects (2.8%) showing bilateral damage and four subjects (3.7%) showing damage only on the left.

The incidence of hearing damage, considering all injuries likely related to occupational exposure, was 35% lower than expected. If only bilateral injuries are considered, the incidence is 70% lower.

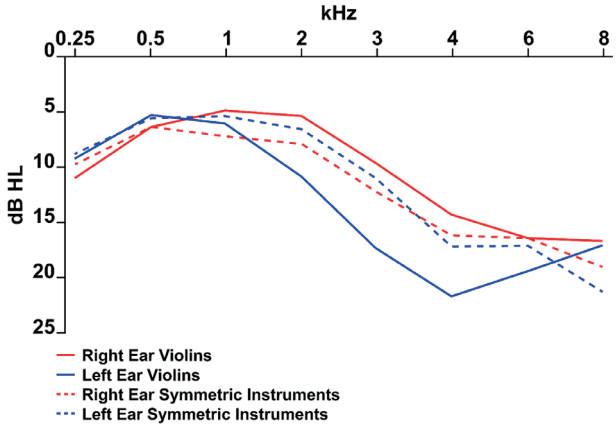

Figure 2. Comparison of average hearing thresholds (right/left) of violinists with those of symmetrical instrument players.

3.4. Longitudinal Assessment

For 80 musicians (61 males and 19 females), audiometric data obtained about 11 years earlier were available, allowing a longitudinal analysis to assess any worsening in hearing thresholds.

The male subjects at the first check-up had an average age of 40 years and an average orchestral length of service of 20.2 years; the females 35.6 and 15.4 years respectively.

The audiometric thresholds between the first examination and the one performed at an average interval of 11 years show minimal worsening bilaterally only at 3, 4, and 8 kHz in males (p-value < 0.01). No differences were observed among females. (Table 3).

The observed audiometric threshold changes are similar to the threshold increases expected from physiological aging [28]. According to the ISO 1999:2015 database A, the hearing threshold at the 50th percentile between ages 40 and 50 at 4 KHz increases by 8 dB HL for males and 5 dB HL for females; in our sample, these increases are never exceeded.

Table 3. Longitudinal audiometric thresholds average dBHL (decibel Hearing Level) of the population by gender.

		Right ear KHz					Left ear KHz				
Subjects	**Average value dBHL**	**2**	**3**	**4**	**6**	**8**	**2**	**3**	**4**	**6**	**8**
Males (61)	*1st evaluation*	9.5	13.3	16.5	21.1	19.6	9.1	12.9	19.4	21.8	20.7
	last evaluation	8.9	15.4	22.1	21.6	25.7	9.1	16.8	24.3	24.5	27
	p-value	ns	**	**	ns	**	ns	**	**	ns	**
Females (19)	*1st evaluation*	9.2	10	12.6	14.5	18.9	8.4	11	13.4	16	14.2
	last evaluation	9.7	13.7	12.6	14.7	16.6	11	13.4	18.4	15.3	16
	p-value	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

** $P < 0.01$.

For individual cases, we considered a worsening due to noise exposure those tracings with an increase of more than 10 dB HL in the average hearing threshold at 2, 3, and 4 kHz after subtracting socio-presbyacusis (OSHA method modified by Brunetti, Menzio, and Morra) [29]. According to this criterion, 7 musicians (8.75%) showed changes likely due to noise: 1 harpist (right side), 4 violinists (left side), 1 horn (left), and 1 flute (left).

4. Discussion

The results of our study are consistent with the most frequently reported data in the literature [30-35], confirming that professional orchestral musicians are at a slightly increased risk of noise-induced hearing loss (NIHL) compared to the general population not occupationally exposed to noise, but at a lower risk than workers exposed to industrial noise of similar intensity and duration. The large sample size and the high percentage of adherence to the study (89.9%) are important indicators for the validity and reliability of the results.

The audiometric tests of NIHL in musicians typically exhibits the classical notch at 3–6 kHz, similar to that seen in occupational settings. However, NIHL in musicians tends to be more asymmetrical, influenced by factors such as the specific instrument played, seating position within the orchestra, and the intermittent nature of the exposure. In musicians playing instruments with high-frequency output, NIHL may also be evident at 8 kHz [18].

Johnson et al. [36] showed higher auditory acuity in orchestral musicians than in non-musicians by testing frequencies from 0.25 to 20 KHz; despite their continuous exposure to music, no significant deterioration in hearing thresholds was observed, nor were there notable gender or interaural differences.

In our study, asymmetrical exposure was particularly evident in violinists, with left ear sound levels reaching 92.2 dBA and right ear levels at 81.4 dBA during individual study, This pattern correlates with a significantly higher risk of left-sided unilateral hearing loss. The resulting hearing loss asymmetry is caused by direct exposure to the sound energy produced by the violin. Asymmetric deafness is also observed in hunters, with (in right-handed people) greater damage on the left side, even though the rifle is resting on the right shoulder. This is because, in this case, the source of the sound from the shot is in the barrel of the rifle, which is oriented to the left, and the right ear is protected by posture (raised shoulder and homolateral tilt of the head).

Suen et al. [37] in a general population of over 6,000 subjects older than of 20 years documented that the prevalence of asymmetrical hearing was higher among men and correlated with age: the left ear has a worse hearing threshold than the right ear even in individuals who have not been exposed to noise, especially at high frequencies. Bidelman [38] suggested a possible neurophysiological basis, proposing that the medial olivocochlear efferent system may be more active in the right ear, offering it greater protection against temporary threshold shifts.

In our data, as we noted, there are a few cases of hearing loss among musicians, almost exclusively

associated with occupational exposure above 87 dB(A).

Even in cases of left unilateral hearing loss, the risk of NIHL remains 35% lower than expected, confirming that the impact of classical music on hearing is less harmful than that observed with noise in industrial contexts. Only three cases of bilateral NIHL have been reported to INAIL (National Institute for Insurance against Accidents at Work) and recognized as occupational disease.

One possible physiological explanation for the lower incidence of NIHL in musicians compared to industrial workers is the acoustic profile of music, which differs significantly from that of industrial noise in terms of temporal structure, spectral complexity, and dynamic range. Music, in fact, spans a broader frequency spectrum, exhibits greater dynamic variability, and often includes periods of rest, thereby allowing partial cochlear recovery. [39]. Moreover, compared to industrial noise, music tends to have fewer impulsive peaks and more predictable temporal and spectral patterns, thereby reducing cochlear overload and potential damage [40]. Furthermore, compared to passive exposure to industrial noise, musical sound is self-generated through intentional motor activity, which may enhance activation of the medial olivocochlear efferent system, which plays a protective role by modulating cochlear amplifier gain and reducing overstimulation of outer hair cells. [38, 41, 42] A further hypothesis could be the influence of neuro-behavioral dynamics, which could explain the lower incidence of hearing damage for the perception of harmonic sounds, recognized and processed at the cortical level, compared to the perception of indistinct and unprocessed noises. The lower incidence of NIHL in musicians compared to industrial workers allows us to advance the hypothesis that it is not only the overall intensity of the noise that determines the damage, but also the composition of more or less harmonic frequencies could play a significant role. It could be hypothesized that certain situations trigger endolymphatic biomechanical alterations, which may be selectively harmful to specific topotonic areas. These aspects, still little known, could be better analyzed in a research project with the study of otoacoustic emissions [43] and with the evaluation of the 'dead regions' using the TEN test" [44].

Kähäri et al. [33] conducted a 16-year follow-up study of hearing thresholds in classical musicians and found no increase in hearing thresholds relative to normative values. This finding is similar to what was shown in our longitudinal evaluation. The reduced progression of hearing impairment over the years suggests that the cochlear auditory damage observed develops mainly during the first years of exposure. It should be kept in mind that in musicians, study can begin as early as 4-7 years of age, and artistic activity, with consequent exposure to acoustic risk, is very intense in the younger years in connection with prolonged study for examinations, competitions, and auditions. Similarly, Behar et al. [34], who measured hearing threshold variations over 5 years in orchestral musicians showed that thresholds did not change over this period, but emphasized that, in the orchestra he studied, measured exposures were usually <85 dB(A) as Leq.

The reduced progression of hearing impairment over the years suggests that the auditory damage of the cochlea observed develops mainly in the first years of exposure. It should be kept in mind that in musicians, study can begin as early as 4-7 years of age and artistic activity, with consequent exposure to acoustic risk, is very intense in the younger years in connection with prolonged study for examinations, competitions and auditions

Recently, the journal La Medicina del Lavoro published an article by F. Tomasina et al. titled Audiometric Database of Academic Musicians in Uruguay [45], which concludes that orchestra musicians, choir singers, and opera soloists experience greater hearing loss than the populations defined by ISO standards. The findings by Tomasina et al. are not comparable to ours for several reasons. The decrease in hearing threshold is only assessed at frequencies of 2000 and 4000 Hz and does not evaluate other important frequencies, particularly 3000 Hz, which is crucial for estimating the progression of noise-induced hearing loss. It is calculated by averaging the hearing thresholds of both ears, whereas in our study, we considered each ear separately because, based on literature and our environmental

measurements, asymmetrical exposure—such as with violins and violas—is common. Additionally, their comparison does not consider damage expected from similar levels of industrial noise exposure. Furthermore, Tomasina's study characterizes the population with heterogeneous exposure, including orchestra musicians, symphonic band members, choir singers, opera soloists, and students. The exposure was also measured as Leq rather than as a weekly exposure estimate, which fails to account for variations in repertoire and exposure duration.

To contribute effectively to knowledge about hearing risks among musicians in opera and symphony orchestras and to propose preventive strategies, researchers must agree on methods to assess exposure, conduct investigations, and evaluate damage.

In this regard, additional studies that better define and identify hearing subjectivity are needed to explore the importance of disorders like tinnitus, hyperacusis, sound distortions, diplacusis, and the sensation of occlusion, which can cause specific "occupational" damage in this group.

5. Conclusion

Our study, in agreement with the majority of studies in the literature, confirms that opera orchestra musicians, although exposed to high noise levels, mostly with Lex-w above 87 dB(A), show a slightly higher risk of hearing damage than the general population, occupationally not exposed to noise, and a significantly lower risk than workers exposed to occupational noise of similar intensity and duration.

It is suggestive that among musicians, who do not usually use hearing protectors, cases of hearing loss are found almost exclusively related to occupational exposure above 87 dB(A).

The longitudinal analysis, comparing audiograms taken more than 10 years apart, shows a progression of hearing loss compatible with physiological ageing.

The high levels of noise impact and related auditory symptoms reported by musicians during rehearsals and performances, and the finding of some cases of hearing loss, however, call for risk reduction interventions:

- Information and training, compulsory under DL 81/2008, emphasizing the relationship between work activity and hearing damage, raising awareness of the use of hearing protectors, particularly during personal practice and rehearsals. In particular, it is of fundamental importance to protect violinists against acoustic trauma in the left ear.
- Exposure reduction should be implemented not through questionable interventions, such as screens and panels (difficult to accommodate in the often-cramped spaces of rehearsal rooms and orchestra pits, which reflect sound and can interfere with musicians' acoustic perception), but by designing the acoustic 'qualification' of spaces. Particular attention must be paid to rehearsal rooms for the containment of sound levels through adequate space for the number of musicians, reduction of sound wave reflections caused mainly by flat walls, and reduction of reverberation through the use of suitable sound-absorbing materials.
- Supplying earplugs, proposing those that are most comfortable and least interfering with musical performance to be used mainly while studying or in some rehearsals.
- Preventive and periodic health surveillance to be carried out with medical examinations and audiological examinations, also in cooperation with audiological specialists as set out in the Guidelines for the Music Sector[6]. It is essential to provide first checks since the beginning of the instrument's study in Conservatories and Academies and to carry out audiological assessments upon employment in orchestras. Subsequent audiological evaluations should be carried out periodically, with more frequent assessment in the first years of activity. To plan a widespread and coordinated prevention activity, it is desirable to discuss the Lyric and Symphonic Foundations, the Lyric Theaters, the Academies, the Philharmonics, the Conservatoires, and other Higher Arts and Music Education (AFAM).

Funding: This research received no external funding. The study was carried out through an agreement between the Teatro alla Scala and the University of Milan. It was approved by the university ethics committee at its meeting on 18/04/2023 (opinion number 48/23).

Informed Consent Statement: Participants were first provided with information on the purpose and methods of the study, as well as on the processing of personal data. Each participant signed both consents, which were collected by the university staff.

Acknowledgments: Mr Fabrizio Lencioni and LEDISO Italia SrL, which provided free loan of the equipment for carrying out the audiological examinations.
Dr Dominique Meyer Superintendent of Teatro alla Scala and the entire management who supported the realization of the study project.
The professors members of the University of Milan ethical committee and of Teatro alla Scala who evaluated the project and allowed to realize it.
Orchestral professors of the Teatro alla Scala for willingly joining the study.
Dr. Pietro Bianchi (Occupational Physician), ms. Lucia Giglio and the nurses of the Teatro alla Scala for organisational support
Medical doctors (Dr Mirko Aldè, Dr Giorgio Lilli) and Audiologist technicians and training students from the Audiology SSD of the Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico Milano, University of Milan, who helped to perform audiological evaluations at the Fondazione Teatro alla Scala.
Dr Antonio Mancuso who read the article and provided some hintful comments.
Dr Luigi Flaminio Ghilardini for his contribution to the creation of the iconographic apparatus.

Declaration of Interest: The authors declare the absence of conflicts of interest.

Author Contribution Statement: TC, AT, OC, UA, EF and FDB contributed to the conception and design of the research; TC, AT, OC, UA and FDB contributed to the analysis of the results. TC and OC wrote the final manuscript with the support of UA and FDB; EF, VC, AT, RB, VT and LM performed the hearing tests;LP designed with CT phonometric investigation; LP performed the phonometric evalutation; LB performed the statistical analytical

Declaration on the use of AI: No chatbot was used.

References

1. Lie A, Skogstad M, Johannessen H A, et al. Occupational noise exposure and hearing: a systematic review. *Int Arch Occup Environ Health*. 2016 Apr; 89(3):351-72. Doi: 10.1007/s00420-015-1083-5. Epub 2015 Aug 7. PMID: 26249711; PMCID: PMC4786595.
2. Hutchinson KM, Alessio HM, Hoppes S, et al. Effects of Cardiovascular Fitness and Muscle Strength on Hearing Sensitivity. *J Strength Cond Res*. 2000;14(3):302-309. Doi: 10.1519/00124278-200008000-00010
3. Zhang X, Ni Y, Liu Y, et al. Screening of noise-induced hearing loss (NIHL)-associated SNPs and the assessment of its genetic susceptibility. *Environ Health*. 2019;18(1):30. Published 2019 Apr 4. Doi: 10.1186 /s12940-019-0471-9
4. Hake R, Kreutz G, Frischen U, et al. A Survey on Hearing Health of Musicians in Professional and Amateur Orchestras. *Trends Hear*. 2024;28:23312165241293762. Doi: 10.1177/23312165241293762
5. European Commission - Directorate General for Employment, Social Affairs and Equal Opportunities - Unit F4 . Chapter 8: The music and entertainment sectors" In: Non-binding guide of good practice for the application of Directive 2003/10/EC [Article in Italian] Luxembourg, 2009 ISBN 978-92-79-11342-0. Doi: 10.2767/21285
6. Permanent Advisory Commission for Occupational Safety and Health "Guidelines for the music and leisure sector, pursuant to Article 198 of Legislative Decree 81/2008, as amended – approved at the meeting of 7 March 2012". [Article in Italian] https://www .lavoro.gov.it/sites/default/files/archivio-doc-pregressi /SicurezzaLavoro/LG_art_198_musica_articolato.pdf
7. Morata TC. Young people: their noise and music exposures and the risk of hearing loss. *Int J Audiol*. 2007;46(3): 111-112. Doi: 10.1080/14992020601103079
8. Moore BCJ, Lowe DA, Cox G. Guidelines for Diagnosing and Quantifying Noise-Induced Hearing Loss. *Trends Hear*. 2022;26:23312165221093156. Doi: 10.1177 /23312165221093156
9. Qian CL, Behar A, Wong W. Noise exposure of musicians of a ballet orchestra. *Noise Health*. 2011;13(50): 59-63. Doi: 10.4103/1463-1741.74002
10. Axelsson A, Lindgren F. Hearing in classical musicians. *Acta Otolaryngol Suppl*. 1981;377:3-74.
11. Westmore GA, Eversden ID. Noise-induced hearing loss and orchestral musicians. *Arch Otolaryngol*. 1981;107(12):761-764. Doi: 10.1001/archotol.1981 .00790480037010
12. Phillips SL, Mace S. Sound level measurements in music practice rooms. Music Performance Research. 2008 Jan;2(1):36-47.
13. Zelený M, Navrátilová A, Kamýcek Z, Vlk Z. [Relation of hearing disorders to the acoustic composition of the

working environment of musicians in a wind orchestra] Vztah poruch sluchu k akustické skladbě pracovního prostredí hudebniků v dechovém orchestru. *Cesk Otolaryngol*. 1975;24(5):295-9. Czech. PMID: 1204072.
14. Laitinen H, Poulsen T. Questionnaire investigation of musicians' use of hearing protectors, self reported hearing disorders, and their experience of their working environment. *Int J Audiol*. 2008;47(4):160-168. Doi: 10.1080/14992020801886770
15. Behar A, Wong W, Kunov H. Risk of hearing loss in orchestra musicians: Review of the literature. *Medical Problems of Performing Artists*. 2006; 21(4):164-8. Doi: 10.21091/mppa.2006.4035.
16. Zbysińska M, Lachowska M. Hearing damage caused by noise in classical musicians. *Pol Otorhino Rev*. 2020;9(2):41-53. Doi: 10.5604/01.3001.0014.1545
17. Behar A, Chasin M, Cheesman M. Noise Control: A Primer. Singular Pub. Group/Thomson Learning. San Diego, 2000 url :lcp:noisecontrolprim0000beha:epub:d3e0e7f9-5d58-4592-b4c5-a5c2e4e74027
18. Chasin M. Musicians and the prevention of hearing loss. Singular Publishing Group Inc. San Diego, CA. (April 1996) ISBN#1565936264.
19. Pyykkö I, Toppila E, Zou J, Erna K. Individual susceptibility to noise-induced hearing loss. *Audiol Med*. 2007;5(1):41-53 ISSN 1651-386X.
20. Couth S, Prendergast G, Guest H, et al. Investigating the effects of noise exposure on self-report, behavioral and electrophysiological indices of hearing damage in musicians with normal audiometric thresholds. *Hear Res*. 2020;395:108021. Doi: 10.1016/j.heares.2020.108021
21. Aazh H, Danesh AA, Moore BCJ. Internal Consistency and Convergent Validity of the Inventory of Hyperacusis Symptoms. *Ear Hear*. 2021;42(4):917-926. Doi: 10.1097/AUD.0000000000000982
22. Couth S, Prendergast G, Guest H, et al. A longitudinal study investigating the effects of noise exposure on behavioural, electrophysiological and self-report measures of hearing in musicians with normal audiometric thresholds. *Hear Res*. 2024;451:109077.
23. Sataloff RT. Hearing loss in musicians. *Am J Otol*. 1991;12(2):122-7.
24. DECRETO LEGISLATIVO 9 aprile 2008, n. 81. Attuazione dell'articolo 1 della legge 3 agosto 2007, n. 123, in materia di tutela della salute e della sicurezza nei luoghi di lavoro. Gazzetta Ufficiale della Repubblica Italiana - Serie Generale n. 101; 30 aprile 2008.
25. Pira E, Bosio D, Merluzzi F. (Prevention of noise-induced hearing damage in the workplace - guidelines proposed by the Italian Society of Occupational Medicine and Industrial Hygiene) La prevenzione dei danni uditivi da rumore in ambiente di lavoro-linee guida proposte dalla Società Italiana di Medicina del Lavoro e Igiene Industriale. [in Italian] Ed. PIME 2003. ISBN:8879631500
26. UNI EN ISO 1999:2015 Acoustics - Estimation of noise-induced hearing loss [Article in Italian].
27. Parati L. Risk Assessment to Surveillance Planning in a Lyric Theatre: a Focus on Noise Exposure in Orchestra and Chorus. Practical Aspects, Problems and Peculiarities. Occupational Voice Symposium, UCL London, 2017.
28. Roth TN, Hanebuth D, Probst R. Prevalence of age-related hearing loss in Europe: a review. *Eur Arch Otorhinolaryngol*. 2011;268(8):1101-1107. Doi: 10.1007/s00405-011-1597-8
29. Brunetti B, Menzio P, Morra B. [Juridical and methodological problems in the evaluation of occupational hearing loss.]. Problemi giuridici e metodologici nella valutazione delle ipoacusie professionali. [Article in Italian]. *Acta Otorhinolaryngol Ital*. 1985;5 Suppl 6 :3-34.
30. Royster JD, Royster LH, Killion MC. Sound exposures and hearing thresholds of symphony orchestra musicians. *J Acoust Soc Am*. 1991;89(6):2793-2803. Doi: 10.1121/1.400719
31. Schmidt JH, Pedersen ER, Juhl PM, et al. Sound exposure of symphony orchestra musicians. *Ann Occup Hyg*. 2011;55(8):893-905. Doi: 10.1093/annhyg/mer055
32. Obeling L, Poulsen T. Hearing ability in Danish symphony orchestra musicians. *Noise Health*. 1999;1(2):43-49.
33. Kähäri KR, Axelsson A, Hellström PA, Zachau G. Hearing development in classical orchestral musicians. A follow-up study. *Scand Audiol*. 2001;30(3):141-149. Doi: 10.1080/010503901316914511
34. Behar A, Chasin M, Mosher S, Abdoli-Eramaki M, Russo FA. Noise exposure and hearing loss in classical orchestra musicians: A five-year follow-up. *Noise Health*. 2018;20(93):42-46. Doi: 10.4103/nah.NAH_39_17
35. Eaton S, Gillis H. Review of Orchestra Musicians' Hearing Loss Risks. *Canadian Acoust*. 2002;30(2):5-12.
36. Johnson DW, Sherman RE, Aldridge J, Lorraine A. Extended high frequency hearing sensitivity. A normative threshold study in musicians. *Ann Otol Rhinol Laryngol*. 1986;95(2 Pt 1):196-202. Doi: 10.1177/000348948609500219
37. Suen JJ, Betz J, Reed NS, Deal JA, Lin FR, Goman AM. Prevalence of Asymmetric Hearing Among Adults in the United States. *Otol Neurotol*. 2021;42(2):e111-e113. Doi: 10.1097/MAO.0000000000002931
38. Bidelman GM, Bhagat SP. Right-ear advantage drives the link between olivocochlear efferent 'antimasking' and speech-in-noise listening benefits. *Neuroreport*. 2015;26(8):483-487. Doi: 10.1097/WNR.0000000000000376
39. Reybrouck M, Podlipniak P, Welch D. Music and Noise: Same or Different? What Our Body Tells Us. *Front Psychol*. 2019;10:1153. Doi: 10.3389/fpsyg.2019.01153
40. Strasser H, Irle H, Legler R. Temporary hearing threshold shifts and restitution after energy-equivalent exposures to industrial noise and classical music. *Noise Health*. 2003;5(20):75-84.

41. Perrot X, Collet L. Function and plasticity of the medial olivocochlear system in musicians: a review. *Hear Res*. 2014;308:27-40. Doi: 10.1016/j.heares.2013.08.010
42. de Boer J, Thornton AR. Neural correlates of perceptual learning in the auditory brainstem: efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *J Neurosci*. 2008;28(19):4929-37. Doi: 10.1523/JNEUROSCI.0902-08.2008. PMID: 18463246]
43. Probst R, Lonsbury-Martin BL, Martin GK. (1991). A review of otoacoustic emissions. *The Journal of the Acoustical Society of America*. 89(5):2027–2067.
44. Moore BCJ, Huss M, Vickers DA, Glasberg BR, AlcántaraJI. A test for the diagnosis of dead regions in the cochlea. *British Journal of Audiology*. 2000;34(4), 205–224.
45. Tomasina F, González AE, Pisani Biriel MA, et al. Audiometric Database of Academic Musicians in Uruguay. *Med Lav*. 2025;116(3):16424. Doi: 10.23749/mdl.v116i3.16424

La Medicina del Lavoro

Original article

Med. Lav. 2025; 116 (5): 17196
DOI: 10.23749/mdl.v116i5.17196

Steady-State Auditory Evoked Potentials in Workers Exposed to Occupational Noise

Mariana Keiko Kamita[1*], Clayton henrique Rocha[1], Liliane Aparecida Fagundes Silva[1], Alessandra Giannella Samelli[1], Carla Gentile Matas[1]

[1]Department of Physiotherapy, Speech-Language and Hearing Sciences and Occupation Therapy, Faculdade de Medicina FMUSP, Universidade de São Paulo, São Paulo, SP, Brazil

Keywords: Hearing; Noise; Noise Effects; Auditory Evoked Potentials; Hearing Tests; Noise, Occupational

Abstract

Background: *The present study aims to investigate neural synchrony, as measured by Auditory Steady State Response (ASSR), in individuals with normal hearing who are exposed and not exposed to occupational noise, thereby providing insights into hidden hearing loss within the central auditory nervous system, and justifying the importance of exploring auditory neural function in populations at risk.* **Methods:** *A cross-sectional study involved 30 noise-exposed individuals in the Study Group and 30 unexposed individuals in the Control Group, all paired by an average age of 35 years. The following procedures were performed on all individuals: clinical and occupational history, meatoscopy, immitanciometry, pure tone audiometry, speech audiometry, and ASSR (40Hz). We analyzed the audiometric hearing thresholds at frequencies of 1 kHz and 4 kHz, the electrophysiological thresholds estimated by ASSR, and the comparison of the differences between them: the thresholds estimated by ASSR and the audiometry thresholds. The data were analyzed using both descriptive and inferential statistics. P-values ≤ 0.05 were considered significant.* **Results:** *When comparing hearing thresholds at 1 kHz and 4 kHz between groups, we found significant differences, with the SG showing higher hearing thresholds than the CG bilaterally. No significant differences were seen in the electrophysiological thresholds estimated by ASSR, nor in the comparison between the ASSR-estimated threshold and the psychoacoustic hearing threshold groups.* **Conclusions:** *The results of this study suggest that workers exposed to occupational noise did not show detectable changes in neural synchrony in the midbrain, thalamus, or primary auditory cortex when compared to individuals without occupational noise exposure.*

1. Introduction

The initial studies on the pathophysiology of noise-induced hearing loss used animal models. They showed that noise exposure mainly causes mechanical damage to inner ear structures, ischemia, and ionic imbalances, with oxidative stress also playing a role. New prevention strategies include antioxidants [1-6].

Later animal studies revealed that noise can induce a temporary threshold shift (TTS) but cause permanent degeneration of presynaptic structures and spiral ganglion neurons. At the same time, hair cells remain intact—indicating synaptopathy [7]. This might involve high- and low-spontaneous-rate auditory fibers connecting with inner hair cells; low-spontaneous-rate fibers are more vulnerable, leading to synapse loss without affecting hearing thresholds,

Received 23.05.2025 – Accepted 25.09.2025

* Corresponding Author: Mariana Kamita; E-mail: marikeikokamita@gmail.com

as high-spontaneous-rate fibers remain functional [2, 4, 8, 9]. Such damage may result in difficulties understanding speech amidst noise, problems perceiving temporal cues, hyperacusis, or tinnitus [2, 4, 6, 9-11]. Still, the pathophysiology of hidden hearing loss (HHL) in humans remains unclear [12]; further research using audiological tests, especially those assessing the central auditory nervous system (CANS), is needed [2-4, 6].

Auditory evoked potentials (AEPs), widely used electrophysiological tools in neuroscience, can evaluate the CANS. Among these, the auditory steady state response (ASSR) is instrumental, as it relies on neural synchrony related to sound localization, pitch, and speech-in-noise perception, and can be recorded at various modulation frequencies to assess temporal processing [13-16]. ASSR helps analyze acoustic transmission and detect neural dyssynchrony, which may result from synaptopathy at the auditory nerve's first synapse or along the pathway, potentially causing dyssynchrony throughout [17].

A previous study found that noise-exposed individuals with normal hearing performed worse on speech perception tests than non-exposed controls, possibly due to noise-induced impairments in neural synchrony not reflected in tonal thresholds, suggesting a link to HHL. Therefore, we hypothesize that workers exposed to occupational noise, without hearing threshold impairment, show altered neural synchrony, which can be demonstrated in ASSR recordings. The objective of this study was to evaluate neural synchrony through ASSR in normal-hearing individuals exposed and not exposed to occupational noise, aiming to gain insights into HHL in the CANS.

2. Methods

This cross-sectional study was approved by the institutional research ethics committee (No. 2.435.259), and the research was conducted in accordance with the Declaration of Helsinki. All subjects included in the study received prior explanations about the research, and after agreeing to the terms, they signed the informed consent form.

The present research is part of a larger study that conducts several peripheral and central audiological assessments in normal-hearing adults exposed to occupational noise, to investigate audiological findings and possible changes resulting from noise exposure.

The study included 60 normal-hearing individuals divided into two groups paired by age: the study group (SG) comprised 30 individuals exposed to occupational noise and with a mean age of 35.60 years, and the control group (CG) comprised 30 individuals not exposed to occupational noise and with a mean age of 35.37 years.

The inclusion criteria for both groups were: male individuals with normal-hearing thresholds bilaterally, absence of earwax and middle ear alterations, aged at least 18 years old and no more than 50 years, with no history of ear diseases or surgeries, no tinnitus, not taking medication or using potentially ototoxic treatments, and no exposure to chemical products. Furthermore, in addition to the aforementioned criteria, the SG required exposure to occupational noise above 85 dB HL for one year or more. In contrast, the CG required the absence of such exposure.

The workers were chosen based on the university's environmental risks prevention program, which outlines the risks each worker faces during their workday. The SG included individuals who worked in maintenance at the university and were exposed to intermittent noise (Lavg [average sound pressure level over a period of time]: 88 dBA; minimum: 75 dBA, maximum: 111 dBA; 69% of the daily dose) during their 8-hour workday for an average of 8.6 years (SD: 6.1 years) in their current position. All of them used Hearing Protection Devices (HPDs) during their shifts, with the plug type being most common (70%). The CG also consisted of workers from the same institution, but from departments without noise exposure, mostly from administrative sectors. The detailed characterization of the sample is shown in Table 1, and the history of occupational and non-occupational noise exposure is shown in Table 2.

Initially, the following procedures were performed: clinical and occupational history; meatoscopy (Mini 3000, Heine); acoustic immittance measurements (AT235, Interacoustics), including tympanometry and the assessment of ipsi- and contralateral acoustic reflexes; and pure tone audiometry (PTA)

Table 1. Characterization of the sample.

Variable		Study Group (n=30)	Control Group (n=30)
Age range (in years)	*Mean*	35.60	35.37
	SD	7.10	7.56
	Minimum	23.00	22.00
	Maximum	50.00	49.00
Educational level, n (%)	*High School*	14 (46.7%)	6 (20.0%)
	Technical Education	3 (10.0%)	1 (3.3%)
	Incomplete Higher Education	7 (23.3%)	2 (6.7%)
	Higher Education Complete	6 (20.0%)	21 (70.0%)
Complaints & otologic history n (%)	*Hyperacusis*	3 (10.0%)	0 (0.0%)
	Itching	3 (10.0%)	5 (17.0%)
	Difficulty listening in noise	3 (10.0%)	0 (0.0%)

Legend: n- Sample number; %- Percentage.

Table 2. History of occupational and non-occupational exposure to noise.

		Study Group (n=30)	Control Group (n=30)
Occupational exposure to noise (in years)	*Mean*	13.86	Not applicable
	SD	8.33	
	Minimum	1.00	
	Maximum	32.00	
Non-occupational exposure to noise, n (%)	*Headphones*	17 (56.7%)	14 (46.7%)
	Stadium/Autodrome	5 (16.7%)	5 (16.7%)
	Church	12 (40%)	7 (23.3%)
	Shows/Parties	5 (16.7%)	13 (43.3%)

Legend: n- Sample number; %- Percentage.

(MA42, MAICO), where hearing thresholds were considered within the normal range of ≤25 dB HL at 250 to 8000 Hz [19].

Once the individuals received normal results in the assessments mentioned above and met the inclusion criteria, they underwent an electrophysiological assessment using the ASSR with the intelligent hearing system (ANSI S3.7-1996) Smart EP model, conducted in an acoustically treated room. The individual was seated comfortably in a reclining chair.

First, the skin was cleaned with abrasive paste at the electrode placement sites, following international standard IES 10-20 [20]: on the vertex (Cz) for the active electrode, on the mastoids of the left (M 1) and right (M 2) ears for the reference electrodes, and on the forehead (Fpz) for the ground electrode. Electrodes were attached using electrolytic paste and microporous tape, ensuring impedance values of less than 5 kOhms. The patient received the acoustic stimulus through ER-3 A insert earphones, stimulating both ears simultaneously.

The ASSR was performed with modulation at 40 Hz, using sinusoidal acoustic stimulation with 100% amplitude modulation and 100% frequency modulation at frequencies of 1 and 4 kHz. Up to 400 stimuli were presented, divided into 20 sweeps of 20 stimuli each, with high-pass and low-pass filters set from 30 to 3000 Hz.

Evaluation began at 1 kHz, followed by 4 kHz. The initial intensity was set at 80 dBnHL, decreasing by 10 dBnHL steps until the electrophysiological threshold was found for each ear, especially if one ear responded and the other did not.

The electrophysiological thresholds (dB SPL) were converted into estimated thresholds (dB HL) based on ISO 389-2, with corrections of 0 dB at 1 kHz and -6 dB at 4 kHz, as used in other studies [21, 22]. The difference between the estimated electrophysiological threshold and the audiometric threshold was then calculated.

Data were analyzed using both descriptive and inferential statistical methods. The Shapiro-Wilk test was used to assess the distribution of the sample, complemented by visual inspection of histograms. Since the data followed a normal distribution pattern, parametric tests were used for further analysis. A repeated measures ANOVA was performed, with the ear as the repeated measure and the group as the between-subjects factor. Variance equality was checked with Levene's test, and sphericity was verified with Mauchly's test. If necessary, the Greenhouse-Geisser correction was applied. Effect sizes were evaluated using eta squared (η^2), following Cohen's (1988) guidelines: $\eta^2 = 0.01$ for a small effect; $\eta^2 = 0.06$ for a medium effect; and $\eta^2 = 0.14$ for a large effect.

3. Results

Initially, a descriptive analysis was carried out of the values obtained in the PTA at 1 and 4 kHz (Table 3) and the estimated thresholds obtained by the ASSR (Table 4).

The comparison analysis showed a statistically significant difference between the groups for the thresholds obtained in the PTA at both 1 kHz and 4 kHz, with a medium effect size in both cases, regardless of which ear was assessed (Table 5; Figure 1). As for the estimated thresholds achieved by ASSR, there were no significant differences either between the ears or between the groups (Table 5).

Regarding the difference between the threshold estimated in the ASSR (Table 4) and the threshold obtained through the ATL (Table 3), there was no statistically significant difference between the groups or between the ears evaluated (Table 6).

4. Discussion

Regarding the data described in scientific literature about HHL in humans, a source of uncertainty relies on indirect electrophysiological measurements, because unlike animal studies, cochlear synaptopathy cannot be directly measured in vivo. Additionally, some authors have suggested that these changes trigger a series of alterations in neural processing in regions posterior to these in the CANS [23]. Therefore, more recent studies have included, among other behavioral tests in noisy environments, measures of function using evoked potentials, mainly auditory brainstem response (ABR), as well as envelope following response (EFR). Currently, audiometry remains the gold standard clinical tool in audiology. However, it is known that this dysfunction can be

Table 3. Descriptive analysis of PTA thresholds for the 1 and 4 kHz frequencies by group and ear.

	Ear	Group	Mean	SD	Minimum	Maximum
1 kHz	Right	Control	7.67	5.68	0	20
		Study	10.33	4.54	0	20
	Left	Control	6.83	5.00	0	20
		Study	10.00	5.09	0	15
4 kHz	Right	Control	9.33	6.40	0	20
		Study	13.00	7.02	0	25
	Left	Control	9.03	6.31	0	20
		Study	12.50	7.63	0	25

Legend: kHz = kilo Hertz.

Table 4. Descriptive analysis of the estimated ASSR thresholds for the 1 and 4 kHz frequencies by group and ear.

	Ear	Group	Mean	Standard deviation	Minimum	Maximum
1 kHz	Right	Control	27.3	7.85	20	40
		Study	26.3	7.65	20	40
	Left	Control	26.0	6.75	20	40
		Study	27.3	10.15	20	60
4 kHz	Right	Control	23.0	13.73	14	54
		Study	25.0	11.55	14	54
	Left	Control	25.7	16.42	14	44
		Study	23.7	10.98	14	44

Legend: kHz- kilo Hertz.

Table 5. Comparison of the thresholds obtained in the PTA and the estimated ASSR threshold at each frequency according to group and ear.

		Factor	Sum of squares	*F*	p	η^2
PTA threshold	1 kHz	Group	255.208	5.904	0.018*	0.078
		Ear	10.208	1.183	0.281	0,003
		Group x Ear	1.875	0.217	0.643	<0.001
	4 kHz	Group	381.633	4.818	0.032*	0.065
		Ear	4.800	0.323	0.572	<0.001
		Group x Ear	0.300	0.020	0.888	<0.001
ASSR threshold	1 kHz	Group	0.833	0.008	0.931	<0.001
		Ear	0.834	0.034	0.854	<0.001
		Group x Ear	40.833	1.682	0.200	0.005
	4 kHz	Group	<0.001	<0.001	>0.999	<0.001
		Ear	13.333	0.270	0.605	<0.001
		Group x Ear	120.000	2.428	0.125	0.006

*Legend: kHz- kilo Hertz; * statistically significant difference;* η2*- Eta squared refers to the size of the effect.*

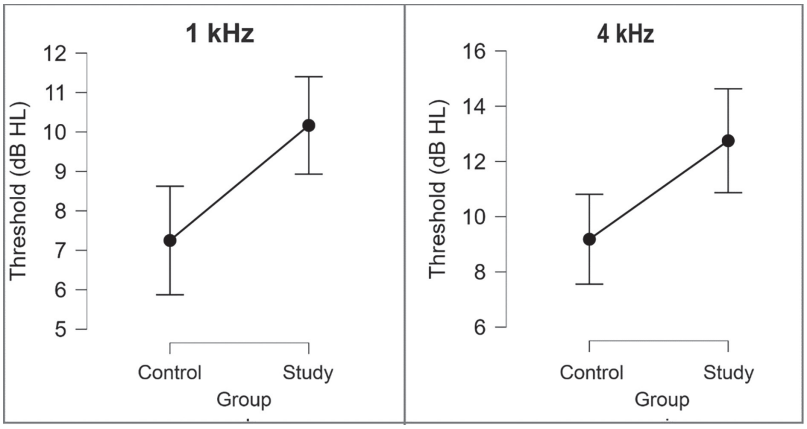

Figure 1. Comparison (in decibel Hearing Level – dB HL) of ATL thresholds between both groups in 1 kHz e 4 kHz (kilo Hertz).

Table 6. Comparison of the difference between the thresholds estimated by ASSR and the PTA threshold for the 1 and 4 kHz frequencies by group and ear.

	Factor	Sum of squares	*F*	*P*	η^2
1 kHz	Group	226,875	2,298	0,135	0,029
	Ear	5,208	0,156	0,694	<0,001
	Group x Ear	25,208	0,757	0,388	0,003
4 kHz	Group	381,633	1,208	0,276	0,017
	Ear	34,133	0,524	0,472	0,002
	Group x Ear	108,300	1,664	0,202	0,005

Legend: kHz- kilo Hertz; η*2- Eta squared refers to the size of the effect.*

"hidden" in the presence of a normal audiogram [12]. Following this trend, and aiming to clarify some of the gaps surrounding the topic, this study aimed to assess neural synchrony through ASSR in normal-hearing individuals exposed and not exposed to occupational noise, in order to identify possible insights into HHL in the CANS, since individuals with central impairment may have a lower agreement between psychoacoustic and electrophysiological estimated thresholds [24].

The study found no significant differences in electrophysiological thresholds between workers exposed and not exposed to occupational noise. However, the exposed group showed higher audiometric thresholds at 1 kHz and 4 kHz compared to the control group. Contrary to the initial hypothesis, which predicted poorer neural synchrony and greater discrepancies between audiometric and electrophysiological thresholds in the exposed group, both thresholds were similar across groups. These findings support what was observed by Grose et al. [25], who conducted ABR, ASSR, and behavioral assessments with speech tests in individuals who frequently attended noisy environments. They found that noise exposure did not lead to changes in any of the behavioral or electrophysiological assessments (ABR and ASSR). The authors emphasized that it was not possible to detect any findings consistent with HHL in their study. However, they suggested that musical experience might have influenced the responses, as many participants in the study group were members of rock bands, and the results could have been affected by the benefits of musical training.

Guest et al. [26], who evaluated individuals with speech perception difficulties in noise using behavioral tests, a questionnaire on previous noise exposure, ABR, and ASSR, observed no correlation between lifetime noise exposure and difficulty listening in noise, as well as the findings of ABR and ASSR. It should be noted that, although HHL likely occurs in humans, no changes in ASSR were observed in the studies mentioned earlier [25,26], including the present one. Guest et al. [26] proposed that potential changes—such as a reduction in the amplitude of wave I of the ABR, which may happen in individuals exposed to noise—do not necessarily result in other detectable alterations along the auditory pathway. Alternatively, ASSR might not be the most sensitive audiological measure for detecting HHL in humans.

On the other hand, Shaheen et al. [27] reported that ASSR has a high sensitivity in detecting HHL in animal models. They assessed distortion product otoacoustic emissions, ABR, and ASSR in mice exposed to noise and noted that only the ASSR test showed altered results, with a reduction in amplitude values. However, it should be noted that the study was conducted on mice living in a controlled environment exposed to noise, and its results were compared with those of mice without noise exposure, who also lived in a controlled environment. This comparison is not possible to carry out on humans.

Furthermore, previous studies have shown that non-human primates are less vulnerable to noise-induced hair cell loss than rodents, although susceptibility to synaptopathy has not yet been investigated [28]. However, this indicates that different species

respond differently to noise exposure and, therefore, more subtle effects of damage to the auditory pathway may not be as easily detected in assessments across all species.

In this way, it is suggested that it is not possible to observe altered ASSR results in humans in the same way as in rodent animals, since occupational and non-occupational noise exposure cannot be measured equivalently. There is also the possibility that humans have auditory system structures that are less susceptible to possible alterations resulting from noise exposure [29].

Despite this, research on humans has yielded findings that differ from those in the present study. Bharadwaj et al. [30], aiming to find an association between hearing assessments and HHL, observed a strong correlation between ASSR and the behavioral performance of individuals exposed to noise with normal hearing. They suggested that the measures obtained by ASSR are sensitive and promising for assessing HHL [30]. Similarly, a study by Mepani et al. [31], involving normal-hearing individuals aged 18 to 63 years with no complaints or history of hearing disorders, found that ASSR has good sensitivity across a range of responses for assessing HHL and that these findings correlate well with behavioral assessments. This suggests that these individuals may have a cochlear or neural deficit and a limited ability to decode words in challenging listening environments.

This difference in findings between ASSR studies may be due to the methods used and the characteristics of the populations studied. Unlike our research, which used an ASSR with a modulation frequency of 40 Hz, Bharadwaj et al. [30] employed a modulation frequency of 100 Hz, and Mepani et al. [31] used frequencies of 128 Hz or 750 Hz, assessing different parts of the CANS.

Along with neural desynchronization, previous research has also identified difficulties in speech understanding in noisy settings within this group. Neural signal transmission from inner hair cells to the auditory nerve occurs across synapses. It is now known that intense noise exposure can selectively harm these synapses, often without a corresponding loss of hair cells. Clinically, this condition exhibits normal audiometric thresholds; however, it presents significant problems in speech perception, particularly in noisy environments. The authors suggest that these issues may be caused by noise exposure, which results in neurodegeneration of spiral ganglion neurons, thereby impairing speech comprehension in such conditions [7, 11, 30, 31].

In the present study, we did not conduct a behavioral assessment, so such a comparison is not possible. However, we observed that 10% of individuals exposed to occupational noise reported difficulty listening in noise, while none of the individuals in the CG did. This suggests that these workers may have some dysfunction in the auditory system. Future studies that incorporate electrophysiological assessments, such as ASSR, alongside behavioral assessments, could help clarify this topic.

Several limitations must be considered in our study, including the lack of detailed information about the duration and level of exposure to extra-occupational noise, as these factors can influence the results for both the control and study groups. Additionally, it is essential to note that the workers participating in this study wore HPDs at work, which may have helped preserve or reduce the risk of noise exposure to the CANS. However, it is known that the use of HPDs is not always practical or consistent across all workplaces or occupational settings (Morata et al, 2024) [32]. Therefore, further research that includes this variable is necessary. Furthermore, the results related to neural synchrony among noise-exposed workers who wear HPDs cannot be generalized to individuals without hearing protection, such as drivers or those exposed to non-occupational noise, nor to those using different types of HPDs. Future studies should examine these specific groups and explore how various types of HPDs might prevent neural synchrony damage, even in cases of lifelong noise exposure. Additionally, the duration of occupational noise exposure should be investigated, as longitudinal studies are essential to assess potential injury to the CANS and its compensatory mechanisms over time.

5. Conclusion

The results obtained in this study suggested that workers exposed to occupational noise did not show detectable changes by ASSR in neural synchrony in

the midbrain, thalamus or primary auditory cortex when compared to individuals without exposure to occupational noise.

Funding: This research was funded by Fapesp, Process no. 2017/25046-4 and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Institution (protocol code 2.435.259 (CAAE: 79905317.7.0000.0065) and date of approval 12/13/2017.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Acknowledgments: None.

Declaration of Interest: The authors declare no conflict of interest.

Author Contribution Statement: M.K.K., C.H.R., data curation, formal analysis, acquisition, writing, review and editing; L.A.F.S., formal analysis, writing, review and editing; A.G.S., conceptualization, data curation, formal analysis, funding acquisition, methodology, project administration, resources, supervision, writing original draft, review and editing; C.G.M., formal analysis, supervision, draft, review and editing. All authors have read and agreed to the published version of the manuscript.

Declaration on the use of AI: None.

References

1. Pienkowski M, Eggermont JJ. Reversible long-term changes in auditory processing in mature auditory cortex in the absence of hearing loss induced by passive, moderate-level sound exposure. *Ear Hear.* 2012;33(3):305-14. Doi: https://doi.org/10.1097/aud.0b013e318241e880
2. Kujawa SG, Liberman MC. Synaptopathy in the noise-exposed and aging cochlea: primary neural degeneration in acquired sensorineural hearing loss. *Hear Res.* 2015;330:191-9. Doi: https://doi.org/10.1016/j.heares.2015.02.009
3. Shi L, Chang Y, Li X, et al. Cochlear Synaptopathy and Noise-Induced Hidden Hearing Loss. *Neural Plast.* 2016;2016:6143164. Doi: https://doi.org/10.1155/2016/6143164
4. Eggermont JJ. Effects of long-term non-traumatic noise exposure on the adult central auditory system. Hearing problems without hearing loss. *Hear Res.* 2017;352:12-22. Doi: https://doi.org/10.1016/j.heares.2016.10.015
5. Falasca V, Greco A, Ralli M. Noise induced hearing loss: The role of oxidative stress. *Otolaryngol Open J.* 2017:S1-5. Doi: http://doi.org/10.1288/00005537-199108000-00005
6. Kobel M, Le Prell CG, Liu J, et al. Noise-induced cochlear synaptopathy: Past findings and future studies. *Hear Res.* 2017;349:148-54. Doi: https://doi.org/10.1016/j.heares.2016.12.008
7. Kujawa SG, Liberman MC. Adding Insult to Injury: Cochlear Nerve Degeneration after "Temporary" Noise-Induced Hearing Loss. *J Neurosci.* 2009;29(45):14077-850. Doi: https://doi.org/10.1523/jneurosci.2845-09.2009
8. Liberman LD, Wang H, Liberman MC. Opposing gradients of ribbon size and AMPA receptor expression underlie sensitivity differences among cochlear-nerve/hair-cell synapses. *J Neurosci.* 2011;31(3):801-8. Doi: https://doi.org/10.1523/jneurosci.3389-10.2011
9. Bharadwaj HM, Verhulst S, Shaheen L, Liberman MC, Shinn-Cunningham BG. Cochlear neuropathy and the coding of supra-threshold sound. *Front Syst Neurosci.* 2014; 21;8:26. Doi: https://doi.org/10.3389/fnsys.2014.00026
10. Knipper M, Van Dijk P, Nunes I, Rüttiger L, Zimmermann U. Advances in the neurobiology of hearing disorders: Recent developments regarding the basis of tinnitus and hyperacusis. *Prog. Neurobiol.* 2013; 111:17–33. Doi: https://doi.org/10.1016/j.pneurobio.2013.08.002
11. Valderrama JT, de la Torre A, McAlpine D. The hunt for hidden hearing loss in humans: From preclinical studies to effective interventions. *Front Neurosci.* 2022;15:16.1000304.Doi:https://doi.org/10.3389/fnins.2022.1000304
12. Grinn SK, Le Prell CG. Evaluation of hidden hearing loss in normal-hearing firearm users. *Front Neurosci.* 2022;16:1005148. Doi: https://doi.org/10.3389/fnins.2022.1005148
13. Gürkan S, Başokçu O, Durankaya SM, et al. Central Auditory Changes Associated with Age-related Hearing Loss. *Clin EEG Neurosci.* 2024;55(4):508-517. Doi: https://doi.org/10.1177/15500594241243116
14. Moore BCJ. The Role of Temporal Fine Structure Processing in Pitch Perception, Masking, and Speech Perception for Normal-Hearing and Hearing-Impaired People. *J Assoc Res Otolaryngol.* 2008;9(4):399-406. Doi: https://doi.org/10.1007/s10162-008-0143-x
15. Plack CJ, Barker D, Prendergast G. Perceptual consequences of "hidden" hearing loss. *Trends Hear.* 2014;18:2331216514550621. Doi: https://doi.org/10.1177/2331216514550621

16. Plack CJ, Leger A, Prendergast G, et al. Towards a diagnostic test for hidden hearing loss. *Trends Hear.* 2016;20:2331216516657466.Doi:https://doi.org/10.1177/2331216516657466
17. Picton TW, John MS, Dimitrijevic A, Purcell D. Human auditory steady-state responses. *Int J Audiol.* 2003;42(4):177-219. Doi: https://doi.org/10.3109/14992020309101316
18. Rocha CH, Lisboa G, Padilha FYOMM, Rabelo CM, Samelli AG. Effects of hearing protector devices on speech intelligibility: the importance of individualized assessment. *Int J Occup Saf Ergon.* 2022;28(2):1227-1234. Doi: https://doi.org/10.1080/10803548.2021.1880763
19. Lloyd LL, Kaplan H. Audiometric interpretation: a manual of basic audiometry. University Park Press: Baltimore. 1978:16-7.
20. Jasper, H.H. The Ten-Twenty Electrode System of the International Federation. *Electroencephalogr Clin Neurophysiol*. (1958)10:371-375.
21. Han D, Mo L, Liu H, et al. Threshold estimation in children using auditory steady-state responses to multiple simultaneous stimuli. ORL *J Otorhinolaryngol Relat Spec.* 2006;68(2):64-8. Doi: https://doi.org/10.1159/000091091
22. Rodrigues GRI, Lewis DR. Potenciais evocados auditivos de estado estável em crianças com perda auditiva cocleares. *Pró fono.* 2010;22(1):37-42. Doi: https://doi.org/10.1590/S0104-56872010000100008
23. Resnik J, Polley DB. Cochlear neural degeneration disrupts hearing in background noise by increasing auditory cortex internal noise. *Neuron*. 2021;17(6):984-996. Doi: https://doi.org/10.1016/j.neuron.2021.01.015
24. Shinn JB, Musiek FE. The Auditory Steady State Response in Individuals with Neurological Insult of the Central Auditory Nervous System. *J Am Acad Audiol.* 2007;18:826-45. Doi: https://doi.org/10.3766/jaaa.18.10.3
25. Grose JH, Buss E, Hall JW 3rd. Loud Music Exposure and Cochlear Synaptopathy in Young Adults: Isolated Auditory Brainstem Response Effects but No Perceptual Consequences. *Trends Hear*. 2017;21:2331216517737417. Doi: https://doi.org/10.1177/2331216517737417
26. Guest H, Munro KJ, Prendergast G, et al. Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure. *Hear Res*. 2018;364:142-151. Doi: https://doi.org/10.1016/j.heares.2018.03.008
27. Shaheen LA, Valero MD, Liberman MC. Towards a Diagnosis of Cochlear Neuropathy with Envelope Following Responses. *J Assoc Res Otolaryngol*. 2015;16(6):727–745. Doi: https://doi.org/10.1007/s10162-015-0539-3
28. Valero MD, Burton JA, Hauser SN, et al. Noise-induced cochlear synaptopathy in rhesus monkeys (Macaca mulatta). *Hear.* 2017;353:213–223. Doi: https://doi.org/10.1016/j.heares.2017.07.003
29. Bramhall NF, Niemczak CE, Kampel SD, et al. Evoked Potentials Reveal Noise Exposure-Related Central Auditory Changes Despite Normal Audiograms. *Am J Audiol.* 2020;29:152-164. Doi: https://doi.org/10.1044/2019_aja-19-00060
30. Bharadwaj HM, Masud S, Mehraei G, et al. Individual differences reveal correlates of hidden hearing deficits. *J Neurosci.* 2015;35(5):2161-72. Doi: https://doi.org/10.1523/jneurosci.3915-14.2015
31. Mepani AM, Verhulst S, Hancock KE, et al. Envelope following responses predict speech-in-noise performance in normal-hearing listeners. *J Neurophysiol.* 2021;125(4):1213-1222. Doi: https://doi.org/10.1152/jn.00620.2020
32. Morata TC, Gong W, Tikka C, Samelli AG, Verbeek JH. Hearing protection field attenuation estimation systems and associated training for reducing workers' exposure to noise. *Cochrane Database Syst Rev.* 2024;5(5):CD015066. Doi: https://doi.org/10.1002/14651858.CD015066.pub2

La Medicina del Lavoro

Original article

Med. Lav. 2025; 116 (6): 17715
DOI: 10.23749/mdl.v116i6.17715

Occupational Risk for Headache Disorders in Female Registered Nurses. A Retrospective Study

Gabriele d'Ettorre[1,*], Roberta Tornese[2], Andrea Baldassarre[3]

[1]Department of Occupational Medicine, Vito Fazzi Hospital, Lecce, Italy
[2]Vito Fazzi Hospital. Local Health Authority, Lecce, Italy
[3]Prevention Department of Occupational Health and Safety Service, Local Health Authority, Lecce, Italy

Keywords: Shift-Work; Work-Related Stress; Healthcare Worker

Abstract

Background: *Prevention of headache disorders (HDs) among healthcare workers in hospital settings remains a challenge for organizations and employees worldwide. The goals of the present retrospective study were both to analyze the 1-year prevalence of any primary HDs among female registered nurses (RNs) employed in hospital settings and to investigate the relationship between occupational risk factors and HDs.* **Methods:** *We analyzed the occupational medicine database of RNs employed in a large hospital. The sample included 975 female RNs; the diagnostic criteria were based on the International Classification of Headache Disorders, 3rd edition (beta version).* **Results:** *One-year prevalence of any HD was 45.9%; tension-type headache (TTH) was the most commonly reported headache type (by 25.6% of participants), followed by migraine (17.5%). No association was found between the different headache types and work schedules; TTH was linked to age ≥40 years (OR=1.91; 95% CI=1.41-2.72), duration of service ≥15 years (OR=1.61; 95% CI=1.24-2.38), and number of night shifts >5 per month (OR=1.71; 95% CI=1.09-2.68). A high level of WRS was a significant predictor of TTH.* **Conclusions:** *We found a link between TTH and modifiable risk factors at both the individual and organizational levels. These findings suggest interventions in occupational settings to minimize the occurrence of TTH among RNs. Policy-makers and employers should implement preventive measures to reduce the incidence of HDs among RNs by minimizing modifiable risk factors associated with increased occupational risk.*

1. Introduction

Headache disorders (HDs) represent a public health problem worldwide as they affect up 90% of people during their lifetime. and are recognized the second cause of disability in all age groups and the first among women under 50 years of age [1, 2]. Globally, HDs affect people of all races, income levels, and geographic areas, and occur more frequently in females than in males [3]. HDs are among the top three most common neurological conditions across most age groups, from age five onward, and remain in the top three until age 80 [4-5]. According to a growing body of evidence, people who suffer from headache-related disorders face health consequences that can lead to impaired quality of life and financial cost due to lost productivity resulting from headache-related absenteeism [6,7]. Several studies revealed increased risk of cardiovascular and cerebrovascular events, myocardial infarction, and stroke in people suffering from HDs [8]; moreover, increased prevalence of anxiety and depression was

Received 13.08.2025 – Accepted 19.11.2025
*Corresponding Author: Gabriele d'Ettorre; E-mail: gabriele.det@libero.it

found in people suffering from migraines compared to healthy individuals. [9]. Given the global burden of HDs, the World Health Organization (WHO) endorsed the Global Campaign to Reduce the Burden of Headache, aimed at minimizing the burden of headache worldwide [10]. In a recent study, Thomas et al. [11] found a relationship between headache-attributed disability and lost productivity in occupational settings; interestingly, the authors demonstrated that investment in structured headache services for treatment of HDs is expected as cost saving besides cost-effective, given that relief of disability through effective treatment of HDs is expected to recover > 20% pro rata of lost productivity.

Many studies performed in workplace settings showed occupational risk factors as triggers for HDs, with females at higher risk than males [1, 12-14]. Work-related stress (WRS) and psychosocial factors present in the workplace, such as low skill discretion, low decision authority, role conflicts, bullying, and effort–reward imbalance, have been found among the triggers of headache, with reciprocal relationships between these factors [15]. Nevertheless, to date, few studies have analyzed the relationship between shift work and HDs, and conclusions are not convergent and suffer from methodological limitations that limit generalizability across occupational settings. In a recent cross-sectional study among nurses, Bjovartn et al. [16] found a relationship between HDs and both shift-work disorder (SWD) and working >20 nights per year, but no relationship with work schedule. Consistent with these findings, a study by Wang et al. [17] found a relationship between HDs and >8 night shifts per month among nursing staff; moreover, seniority of >5 years was found to be a risk factor for HDs. A recent metanalysis [18] focused on seven cross-sectional studies in different occupational settings, showed that individuals working night shifts had a 44% higher risk of developing headaches (HR = 1.44, 95% CI: 1.09-1.90, P = 0.011); furthermore, shift work was found to be associated with a higher incidence of migraines (HR = 1.63, 95% CI: 1.27-2.08, P < 0.001) and night shift work was associated with a decreased incidence of migraines (HR = 0.74, 95% CI: 0.57-0.96, P = 0.024); although the cross-sectional design of such seven checked studies included in the metanalysis did not allow the authors to draw strong conclusions, the findings confirmed the need for further studies on the matter.

Given the current concern regarding the occurrence of HDs in nursing staff working in hospital settings we performed a retrospective survey aimed to analyze the 1-year prevalence of any primary HDs among female registered nurses (RNs) in a large Hospital in Salento, Italy, and to investigate the relationship between occupational risk factors and HDs. The diagnostic criteria for HDs were based on the International Classification of Headache Disorders, 3rd edition (beta version) [19].

2. Methods

We conducted a retrospective survey by analyzing the occupational medicine database of RNs employed in a large hospital (Vito Fazzi Hospital, Lecce) located in Salento (the Southern part of the Puglia region, in Italy), who underwent the routine annual mandatory occupational health surveillance from March 1st, 2024, to February 28th, 2025. The sample included female RNs employed in hospital wards. To evaluate the prevalence of HDs, all RNs were interviewed by the occupational physician during the mandatory health surveillance. The diagnostic criteria were based on the International Classification of Headache Disorders, 3rd edition (beta version) (ICHD-3-beta) [19]. The occupational physician first interviewed the RNs, and the following question was asked: 'Have you suffered from a headache during the last year?' Only RNs who answered "yes" were asked to respond to the other headache items. The screening-positive RNs were asked to report frequency, attack duration, intensity, and accompanying headache symptoms to classify migraine, chronic headache (CH), medication overuse headache (MOH), and tension-type headache (TTH).

Respondents reporting headaches lasting more than 4 h per day on 15 or more days per month were given the label of chronic headache and questioned on medication usage to identify MOH [13]. Trigeminal autonomic cephalalgia, other primary headaches, and secondary headaches were not included in this study. MOH was defined as a chronic headache disorder in which the headache occurs

on 15 or more days per month due to regular overuse of medication; these headaches must have been present for more than 3 months last year [20]. We excluded RNs who had suffered head injuries, road accidents in the past year, or who were professionally exposed to chemicals that can cause HDs. All RNs were screened for SWD using three questions from the International Classification of Sleep Disorders (ICSD) [21]. The questions were: (a) Do you experience either difficulties sleeping or excessive sleepiness? (yes/ no), (b) Is the sleep or sleepiness problem related to a work schedule that makes you work when you usually would sleep? (yes/no), (c) Have you had this sleep or sleepiness problem related to the work schedule for at least 1 month? (yes/no). Participants were identified as suffering from SWD if they answered "yes" to all three questions. The WRS was evaluated according to the INAIL methodology [22, 23], and each RN was assigned a WRS level (high, medium, or low).

Data were analyzed using SPSS (Statistical Package for the Social Sciences) version 14.0. Analysis of the frequency of individual variables was conducted using descriptive statistics. Comparisons between groups were performed using the Mann-Whitney U test for nonparametric data when the groups were independent. The statistical significance was set at $p < 0.05$ for all analyses. The statistical analysis included an adjusted (age, marital status, and children living at home as covariates) logistic regression to calculate the odds ratio (OR) with a 95% confidence interval. In this study, the independent variables were age, length of service, work schedule, number of night shifts per month, SWD, and the dependent variables were the headache types (migraine, TTH, CH, MOH).

3. Results

The study involved 975 female RNs, with a mean age of 48.7 years (SD±9.4); demographic data are shown in Table 1.

The one-year prevalence of any HD among RNs was 448 (45.9%), with TTH being the most commonly reported headache type (by 25.6% of participants), followed by migraine (17.5%). Chronic headache and MOH were reported by 1.7% and 1.1% of RNs, respectively. Most of the RNs were exposed to high WRS (62,7%), and the 1-year prevalence of SWD was 24.4 %.

We found no significant differences in the prevalence of headache, migraine, TTH, chronic headache, or MOH across different work schedules (Table 2).

Similarly, logistic regression analyses adjusted for age, marital status, and children living at home showed no association between the various headache types and work schedule (Table 3).

Table 1. Main characteristics of the study population (female registered nurses).

Variables	n. (%)
Age (years)	
<40	355 (36.4)
≥40	620 (63.6)
Lenght of service (years)	
<15	381 (39.1)
≥15	594 (60.9)
Work schedule	
Day only	76 (7.7)
Two shift rotation	308 (31.6)
Three shift rotation	591 (60.7)
Number of night shifts per month	
0	385 (39.5)
1-5	485 (49.7)
>5	105 (10.8)
Shift work disorder	
No	737 (75.6)
Yes	238 (24.4)
Prevalence of primary headache last year	
Migraine	171 (17.5)
Tension-type headache (TTH)	249 (25,6)
Chronic headache (CH)	17 (1.7)
Medication overuse headache (MOH)	11 (1.1)
Work-related stress	
Low	62 (6.3)
Medium	302 (31)
High	611 (62.7)

Table 2. Prevalence of different headache types.

Variables	Migraine n (%)	TTH n (%)	CH n (%)	MOH n (%)
Age (years)				
<40	80 (22.5)	64 (18)	6 (1.69)	4 (1.12)
≥40	91 (14.7)	185 (29.8)	11 (1.77)	7 (1.13)
Lenght of service (years)				
<15	68 (17.8)	76 (19.9)	7 (1.84)	4 (1.05)
≥15	103 (17.3)	173 (29.1)	10 (1.68)	7 (1.18)
Work schedule				
Day only	14 (18.4)	18 (23.7)	1/(1.32)	1 (1.31)
Two shift rotation	52 (16.9)	70 (22.7)	6 (1.95)	4 (1.30)
Three shift rotation	105 (17.8)	161 (27.2)	11 (1.86)	6 (1.01)
Number of night shifts per month				
0	68 (17.7)	96 (24.9)	7 (1.81)	4 (1.04)
1-5	85 (17.5)	115 (24.5)	8 (1.65)	5 (1.03)
>5	18 (17.1)	38 (32.4)	2 (1.9)	2 (1.9)
Shift-work disorder				
No	129 (17.5)	171 (23.2)	13 (1.76)	6 (0.81)
Yes	42 (17.6)	78 (32.8)	4 (1.68)	5 (2.1)
Work-related stress				
Low	10 (16.1)	8 (12.9)	1 (1.6)	1 (1.6)
Medium	53 (17.5)	70 (23.2)	5 (1.7)	3 (1)
High	108 (17.7)	171 (27.9)	11 (1.8)	7 (1.1)

In the adjusted logistic regression analysis, TTH was associated with age ≥40 years (OR=1.91; 95% CI=1.41-2.72), length of service ≥15 years (OR=1.61; 95% CI=1.24-2.38), and having more than 5-night shifts per month (OR=1.71; 95% CI=1.09-2.68) (Table 3).

The prevalence of TTH was higher among RNs with SWD compared to nurses without SWD (Table 2); adjusted logistic regression analysis showed that TTH was associated with SWD (OR=1.59; 95% CI=1.18-2.21) (Table 3). Migraine, CH, and MOH were not linked to SWD. A high level of WRS was a predictive factor for TTH in RNs (OR=2.64; 95% CI=1.31-5.65), but not for other types of HDs.

After adjusting for age, we found a higher prevalence of TTH among RNs over 40 exposed to high-level WRS than among younger RNs (OR=3.19; 95% CI=2.10-4.85) (Table 4).

Migraine was most common among RNs under 40 years old (OR=3.19; 95% CI= 0.41-0.84), but no significant link was found with length of service, work schedule, night shifts, and SWD; no connection was found between both MOH and CH and the dependent variables examined in the study (Table 3).

4. Discussion

Our study examined the one-year prevalence of any primary HDs in female RNs working in hospital settings. Consistent with other studies [17,24], we found that the most common HDs were TTH and migraine, with prevalence rates aligning with the 1-year global prevalence in

Table 3. Adjusted regression analysis with different headache types as dependent variables.

Variables	Migraine OR (95% CI)*	TTH OR (95% CI)*	CH OR (95% CI)*	MOH OR (95% CI)*
Age (years)				
<40	1	1	1	1
≥40	0.58 (0.41-0.84)	1.91 (1.41-2.72)	1.1 (0.48-2.75)	1.15 (0.31-3,49)
Lenght of service (years)				
<15	1	1	1	1
≥15	0.98 (0.71-1.42)	1.61 (1.24-2.38)	0.92 (0.31-2.48)	1.08 (0.27-3.88)
Work schedule				
Day only	1	1	1	1
Two shift rotation	0.91 (0.49-1.67)	0.94 (0.51-1.68)	1.49 (0.18-12.56)	0.99 (0.11-8.96)
Three shift rotation	0.98 (0.52-1.85)	1.28 (0.71-2.24)	1.42 (0.18-11.17)	0.77 (0.10-6.48)
Number of night shifts per month				
0	1	1	1	1
1-5	0.99 (0.68-1.38)	0.98 (0-72-1.35)	0.91 (0.32-2.52)	0.99 (0.26-3.72)
>5	0.95 (0.51-1.65)	1.71 (1.09-2.68)	1.05 (0.21-5.12)	1.85 (0.33-10.24)
Shift-work disorder				
No	1	1	1	1
Yes	1.01 (0.69-1.49)	1.59 (1.18-2.21)	0.98 (0.35-2.98)	2.61 (0.79-8.65)
Work-related stress				
Low	1	1	1	1
Medium	1.11 (0.53-2.32)	1.91 (0.91-4.46)	1.03 (0.12-8.95)	0.61 (0.06-5.98)
High	1.12 (0.57-2.61)	2.64 (1.31-5.65)	1.12 (0.14-8.91)	0.71 (0.10-5.81)

**Logistic regression analyses with independent variables adjusted for age, marital status and children at home.*

Table 4. Adjusted regression analysis with tension-type headache (TTH) as dependent variable among RNs exposed to high WRS and split for age.

Age (years)	TTH OR (95% CI)*
< 40	1
> 40	3.19 (2.10-4.85)

**Logistic regression analyses with each independent variable adjusted for age, marital status and children living at home.*

the general population reported in the review by Stovner et al. [1].

Interestingly, our research revealed that the prevalence of migraine decreased in individuals over 40 years old, while the prevalence of TTH continued to increase in the same age group; these findings contrasted with the Global Burden of Disease study [25], which found that the prevalence of TTH declines with increasing age, including beyond 65 years, after peaking between ages 35–39. However, consistent with our findings, the review by Onan et al. [26] showed a negative relationship between age and migraine. We hypothesize that occupational exposure to WRS could lead to a higher risk of TTH in older RNs, as evidence suggests that vulnerability to stress increases with age [27].

To date, psychological WRS is widely recognized as a contributing factor to TTH; indeed, while many factors have been reported as headache triggers, stress is by far the most common [28]. The pathway through which WRS leads to TTH is not clearly

understood. Several past studies have highlighted that psychosocial factors appear to play a significant role in the onset of headaches and have shown that excessive psychosocial burdens resulting from work demands, insufficient control over work, and dissatisfaction with uninteresting tasks are associated with HDs [14, 15]. Low skill discretion and low decision authority, role conflict, poor social climate, bullying/harassment, and effort-reward imbalance have consistently been linked to higher odds of headaches in multiple studies examining occupational settings [29, 30].

Regarding the relationship between TTH and SWD, our findings are consistent with the study conducted by Bjorvatn [16], which revealed SWD as a risk factor for HDs, including TTH. The cross-sectional design of our study does not allow us to draw conclusions about the causal relationship between SWD and HDs but emphasizes the need for further investigation into the biological pathway linking SWD and TTH.

To date, although HDs have been suggested as a possible predisposing and sustaining factor of SWD, the causal relationship between SWD and HDs remains unclear, given the potential for a two-way relationship [31] as suggested by some studies that examined the issue and hypothesized that the connection between SWD and HDs is bidirectional [32]. In a recent study, Petit et al. [33] argued that SWD and HDs share a common metabolic cause, as glycogen metabolism has been shown to play a crucial role in both disorders; specifically, sleep disturbances impair glycogen metabolism, leading to disruptions in synaptic function and network plasticity, as observed in HDs. Given the increasing evidence of interrelations between SWD and HDs, interventions focused on sleep hygiene could serve as a strategic approach to prevent HDs.

Interestingly, in our study, we found a connection between TTH and modifiable risk factors at both the individual level (i.e., suffering from SWD) and the organizational level (i.e., working night shifts of more than 5 per month, occupational exposure to high WRS). These findings suggest implementing interventions in occupational settings aimed at reducing the occurrence of TTH in RNs, focusing on minimizing WRS and limiting night shifts to 5 per month.

Regarding the increase in TTH occurrence among RNs suffering from SWD, the findings of our study suggest the need for interventions aimed at preventing SWD. To date, a body of evidence indicates that organizational and individual measures are effective in reducing the impact of shift work on workers' health and in preventing the misalignment between sleep-wake rhythm and shift work that leads to SWD [34-36].

This study had some limitations; firstly, the cross-sectional design does not allow for determining the causal relationship between variables; cross-sectional studies can reveal associations but cannot indicate whether the associated factor is a cause or a consequence or whether there is reciprocity between the variables. Therefore, in future studies, causal relationships among variables should be analyzed using a longitudinal study design.

Secondly, the findings relate to hospital settings and may have been affected by organizational factors specific to the Italian occupational context, and therefore, might not be generalizable to all healthcare environments.

Third, our study did not investigate the "healthy night worker effect." Since vulnerable RNs might have left early in their careers, the risk of SWD among shift work nurses could be underestimated.

Finally, we didn't assess the individual chronotype of the RNs.

5. Conclusion

The knowledge and, therefore, prevention of the professional risk factors triggering HDs can reduce the frequency of the phenomenon and prevent its chronicity, thereby promoting RNs' health and, consequently, their wellness.

Based on our study findings, policy-makers and employers should take preventive actions to lower the incidence of HDs among RNs by reducing modifiable risk factors linked to higher occupational risk.

Funding: This research received no external funding.

Institutional Review Board Statement: This study reports, in an anonymous manner, the results of institutional mandatory occupational health surveillance conducted at

the hospital level and therefore did not require ethical committee approval.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Declaration of Interest: The authors declare no conflict of interest.

Author Contribution Statement: All authors contributed equally to this manuscript.

Declaration on the use of AI: None.

References

1. Stovner LJ, Hagen K, Linde M, Steiner TJ. The global prevalence of headache: an update, with analysis of the influences of methodological factors on prevalence estimates. *J Headache Pain*. 2022;23(1):34. Doi: 10.1186/s10194-022-01402-2
2. Steiner TJ, Husøy A, Stovner LJ. GBD2021: headache disorders and global lost health - a focus on children, and a view forward. *J Headache Pain*. 2024;25(1):91. Doi: 10.1186/s10194-024-01795-2
3. Leonardi M, Martelletti P, Burstein R, et al. The World Health Organization Intersectoral Global Action Plan on Epilepsy and Other Neurological Disorders and the headache revolution: from headache burden to a global action plan for headache disorders. *J Headache Pain*. 2024;25(1):4. Doi: 10.1186/s10194-023-01700-3
4. World Health Organization. Migraine and other headache disorders. 6 March 2024. https://www.google.it/url?esrc=s&q=&rct=j&sa=U&url=https://www.who.int/news-room/fact-sheets/detail/headache-disorders&ved=2ahUKEwjjw_ai1oqHAxWawQIHHUklDQkQFnoECAYQAg&usg=AOvVaw1RzqOLrYF7cAldh1DIwoCj (Last Accessed on 08/08/2025).
5. GBD 2021 Nervous System Disorders Collaborators. Global, regional, and national burden of disorders affecting the nervous system, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet Neurol*. 2024;23(4):344-381. Doi: 10.1016/S1474-4422(24)00038-3
6. Kim BK, Cho SJ, Kim CS, et al. Disability and Economic Loss Caused by Headache among Information Technology Workers in Korea. *J Clin Neurol*. 2021;17:546–557.
7. Di Prinzio RR, Arnesano G, Meraglia I, Magnavita N. Headache in Workers: A Matched Case-Control Study. *Eur J Investig Health Psychol Educ*. 2022;12(12):1852-1866. Doi: 10.3390/ejihpe12120130
8. Mahmoud, AN, Mentias A, Elgendy AY, et al. Migraine and the risk of cardiovascular and cerebrovascular events: A meta-analysis of 16 cohort studies including 1 152 407 subjects. *BMJ Open.* 2018;8:e020498.
9. Irimia P, Garrido-Cumbrera M, Santos-Lasaosa S, et al. Impact of monthly headache days on anxiety, depression and disability in migraine patients: Results from the Spanish Atlas. *Sci Rep*. 2021;11:8286.
10. Saylor D, Steiner TJ. The Global Burden of Headache. *Semin Neurol*. 2018;38(2):182-190. Doi: 10.1055/s-0038-1646946. Epub 2018 May 23.
11. Thomas H, Kothari SF, Husøy A, et al. The relationship between headache-attributed disability and lost productivity: 2. Empirical evidence from population-based studies in nine disparate countries. *J Headache Pain*. 2021;22(1):153. Doi: 10.1186/s10194-021-01362-z
12. Wang Y, Xie J, Yang F, et al. The prevalence of primary headache disorders and their associated factors among nursing staff in North China. *J Headache Pain*. 2015;16:4.
13. Vicente Herrero MT, Ramírez Iñiguez de la Torre MV, Reinoso Barbero, L, et al. Preventive Aspects for Migraine and the Workplace: A European Survey. *Arch Prev Riesgos Labor*. 2021,24:20–33.
14. Magnavita N. Headache in the Workplace: Analysis of Factors Influencing Headaches in Terms of Productivity and Health. *Int J Environ Res Public Health*. 2022;19(6):3712.
15. Urhammer C, Grynderup MB, Appel AM, et al. The Effect of Psychosocial Work Factors on Headache: Results from the PRISME Cohort Study. *J Occup Environ Med*. 2020;62:e636–e643.
16. Bjorvatn B, Pallesen S, Moen BE, Waage S, Kristoffersen ES. Migraine, tension-type headache and medication-overuse headache in a large population of shift working nurses: a cross-sectional study in Norway. *BMJ Open*. 2018;8(11):e022403.
17. Wang Y, Xie J, Yang F, et al. The prevalence of primary headache disorders and their associated factors among nursing staff in North China. *J Headache Pain*. 2015(13);16:4.
18. Wang Y, Xie J, Yang F, et al. The prevalence of primary headache disorders and their associated factors among nursing staff in North China. *J Headache Pain*. 2015(13);16:4.
19. Headache Classification Committee of the International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition (beta version). *Cephalalgia*. 2013;33(9):629-808.
20. Cheung V, Amoozegar F, Dilli E. Medication overuse headache. *Curr Neurol Neurosci Rep*. 2015;15(1):509.
21. American Academy of Sleep Medicine. The international classification of sleep disorders. In: Diagnostic and coding manual. 2nd edn. Westchester, IL, 2005.
22. Di Tecco C, Ronchetti M, Ghelli M, Russo S, Persechino B, Iavicoli S. Do Italian Companies Manage Work-Related Stress Effectively? A Process Evaluation

in Implementing the INAIL Methodology. *Biomed Res Int*. 2015;2015:197156.
23. Persechino B, Valenti A, Ronchetti M, Rondinone BM, Di Tecco C, Vitali S, Iavicoli S. Work-related stress risk assessment in Italy: a methodological proposal adapted to regulatory guidelines. *Saf Health Work*. 2013 Jun;4(2):95-9. Doi: 10.1016/j.shaw.2013.05.002
24. Kuate Tegueu C, Dzudie Tamdja A, Kom F, et al. Headache in the adult population of Cameroon: prevalence estimates and demographic associations from a cross-sectional nationwide population-based study. *J Headache Pain*. 2024;25(1):42.
25. Lancet Neurol (2018) Global, regional, and national burden of migraine and tension-type headache, 1990 –2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol*. 17(11):954–976.
26. Onan D, Younis S, Wellsgatnik WD, et al. Debate: differences and similarities between tension-type headache and migraine. *J Headache Pain*. 2023;24(1):92.
27. Sannes AC, Christensen JO, Nielsen MB, Gjerstad J. Stress-induced headache in the general working population is moderated by the NRCAM rs2300043 genotype. *Scand J Pain*. 2022;23(2):326-332.
28. Cathcart S, Winefield AH, Lushington K, Rolan P. Stress and tension-type headache mechanisms. *Cephalalgia*. 2010;30(10):1250-67.
29. Tynes T, Johannessen HA, Sterud T. Work-related psychosocial and organizational risk factors for headache: A 3-year follow-up study of the general working population in Norway. *J Occup Environ Med*. 2013; 55: 1436–1442.
30. Kim HR. Associations Between Workplace Violence; Mental Health; and Physical Health among Korean Workers: The Fifth Korean Working Conditions Survey. *Workplace Health Saf*. 2021;70:161–172.
31. Yi T, Gao P, Zhu T, Yin H, Jin S. Glymphatic System Dysfunction: A Novel Mediator of Sleep Disorders and Headaches. *Front Neurol*. 2022(19);13:885020.
32. Ferini-Strambi L, Galbiati A, Combi R. Sleep disorder-related headaches. *Neurol Sci*. 2019; 40:107–13.
33. Petit JM, Eren-Koçak E, Karatas H, Magistretti P, Dalkara T. Brain glycogen metabolism: a possible link between sleep disturbances, headache and depression. *Sleep Med Rev*. 2021;(59):101449.
34. Rosa D, Terzoni S, Dellafiore F, Destrebecq A. Systematic review of shift work and nurses' health. *Occup Med* (Lond). 2019;69(4):237-243.
35. d'Ettorre G, Pellicani V. Preventing Shift Work Disorder in Shift Health-care Workers. *Saf Health Work*. 2020;11(2):244-247.
36. Okechukwu CE, Griffiths MD, Carta MG, et al. Biological and practical considerations regarding circadian rhythm and mental health relationships among nurses working night shifts: a narrative review and recommendations. *Riv Psichiatr*. 2022;57(2):67-79.

La Medicina del Lavoro

Original article

Med. Lav. 2025; 116 (6): 16988
DOI: 10.23749/mdl.v116i6.16988

Assessing Cumulative Musculoskeletal Strain in Automotive Mechanics: Insights from Real-World Occupational Analysis

Hadi Daneshmandi[1,*], Reyhane Sadeghian[2,1], Niloofar Rahgoshay[2], Nahideh Sartavi[2], Fardin Negahdari[2], Laleh Golkari[3]

[1]Research Center for Health Sciences, Institute of Health, Shiraz University of Medical Sciences, Shiraz, Iran
[2]Student Research Committee, Shiraz University of Medical Sciences, Shiraz, Iran
[3]Food Industries Educational Group, District 3, The Education Department, Shiraz, Iran

Keywords: Cumulative Load; Automotive Mechanics; Musculoskeletal Disorders

Abstract

Background: *This cross-sectional study aims to assess cumulative loads affecting the lower back, shoulders, and distal upper extremities among automotive mechanics.* **Methods:** *The survey was conducted in automotive repair workshops in Shiraz, involving 157 independent mechanics selected through convenience sampling. Data were collected using a multiple-questionnaire including the Persian Cornell Musculoskeletal Discomfort Questionnaire (P-CMDQ), the Lifting Fatigue Failure Tool (LiFFT), the Shoulder Work Assessment Tool (SWAT), and the Distal Upper Extremity Tool (DUET). Descriptive statistics were used to assess musculoskeletal discomfort, and Partial correlation analyses, adjusted for age and Body Mass Index (BMI), examined the relationships between risk levels from LiFFT, SWAT, and DUET and discomfort reported in the P-CMDQ.* **Results:** *The results showed a high level of musculoskeletal discomfort, especially in the lower back, shoulders, and hands. Risk assessments indicated that the cumulative loads are in the high range for the lower back in 42.7% of cases, the shoulders in 40.8%, and the distal upper extremities in 36.3%. A strong correlation was observed between cumulative load on the lower back and perceived discomfort in this region (r = 0.730), whereas the correlations for the shoulders (r = 0.611) and distal upper extremities (r = 0.537) were moderate.* **Conclusions:** *The findings highlight the significant influence of workplace factors on the musculoskeletal health of automotive mechanics, emphasizing the importance of preventive measures and ergonomic solutions to enhance their health and productivity.*

1. Introduction

Modern industrialization has transformed how human needs are met through the widespread use of advanced machinery, equipment, and complex processes. Although this progress has led to significant economic and technological improvements, it has also posed considerable risks to workers. Industrial environments expose workers to a range of hazards, including physical, chemical, biological, mechanical, psychological, and ergonomic factors, all of which can significantly impact their health, safety, and overall well-being [1, 2]. Among these risks, work-related musculoskeletal disorders (WMSDs) have become a widespread problem across multiple industries, especially in the automotive sector [3, 4].

Musculoskeletal disorders (MSDs) encompass a broad range of conditions that affect the muscles,

Received 20.03.2025 – Accepted 06.08.2025
*Corresponding Author: Reyhane Sadeghian; E-mail: sadeghianreyhane1@gmail.com

joints, tendons, nerves, and bones, and may occasionally involve the circulatory system [5]. MSDs are marked by symptoms such as discomfort, numbness, pain, and limited mobility in the affected areas [6]. The severity of these conditions can vary significantly, ranging from mild, localized discomfort to severe injuries that require medical treatment and extended sick leave [7]. Common examples include lower back pain, neck strain, and carpal tunnel syndrome, all of which can significantly hinder a person's ability to perform daily activities and stay productive [8].

Research on WMSDs has been conducted across various industries, including studies involving hospital staff [9], dentists [10], and office workers [11]. However, studies specifically focusing on WMSDs in the automotive repair and maintenance sector have been limited [12]. In the automotive field, physical work often involves repetitive movements, lifting heavy objects, maintaining poor postures for extended periods, and performing tasks such as repetitive turns, prolonged bending, or excessive leaning [13, 14]. These activities lead to a high rate of MSDs, especially in the lower back, shoulders, wrists, and neck among vehicle mechanics [4]. Consequently, vehicle repair work is consistently ranked as one of the highest-risk jobs for WMSDs, with prevalence rates highlighting the urgent need for ergonomic improvements and better workplace practices [13, 15, 16].

Recent research underscores the high prevalence of WMSDs among workers in the automotive industry [12, 17, 18]. Zhang et al. (2023) found that 32% of automobile maintenance workers in their epidemiological study. 8% of workers experienced WMSDs. The most affected areas were the lower back (17.1%), neck (16.3%), and shoulders (14.5%) [12]. Likewise, Patel et al. (2023) reported that nearly 80% of car garage workers experienced work-related musculoskeletal pain, with the lower back being the most commonly affected site. The shoulder and neck were the second and third-most-affected regions, respectively. Additionally, many workers reported pain in multiple body parts, highlighting how widespread the problem is [19]. Further evidence comes from He et al. (2023), who conducted a systematic review and meta-analysis on the prevalence of WMSDs among workers in China's automobile manufacturing industry. Their findings revealed an overall prevalence of WMSDs of 53.1% (95% Confidence Interval [CI] = 46.3% to 59.9%), with the lower back and waist being the most affected areas (36.5%, 95% CI = 28.5% to 44.5%) [17].

Cumulative loads on the body are crucial in the development of MSDs. These loads result from repeated exposure to physical stressors, such as lifting, carrying, repetitive motions, and holding awkward postures, over time. Extended exposure to these stressors can cause tissue fatigue, microtrauma, and eventually chronic pain or injury in vulnerable areas such as the lower back, shoulders, and distal upper extremities. Assessing cumulative loads is crucial for understanding the long-term effects of work activities on the body and for developing effective interventions [6, 20, 21]. Tools like the Lifting Fatigue Failure Tool (LiFFT), Shoulder Work Assessment Tool (SWAT), and Distal Upper Extremity Tool (DUET) help evaluate these loads and identify high-risk tasks. By measuring cumulative exposure, employers can apply targeted ergonomic solutions to lower the risk of MSDs and support long-term musculoskeletal health [22-24]. These measures not only offset initial costs but also help reduce workers' compensation claims and healthcare expenses, benefiting both employees and the organization. Therefore, this study aimed to evaluate cumulative injuries impacting the lower back, shoulders, and distal upper extremities among automotive mechanics.

2. Methods

2.1. Study Design, Setting, and Population

This cross-sectional study was carried out in automotive repair workshops in Shiraz, focusing on automotive mechanics as the research population. The inclusion criteria for this study included male gender, willingness to participate, at least one year of work experience, no involvement in secondary employment, no history of musculoskeletal disorders (either chronic or acute) in any body region, no prior musculoskeletal surgeries, no use of medications related to musculoskeletal conditions, and

no use of protective equipment aimed at reducing musculoskeletal disorders. The exclusion criterion was unwillingness to continue participation during the study. A total of 157 automotive mechanics were included in the study through convenience sampling. All participants were self-employed mechanics working in independent repair settings. The study received approval from the Ethics Committee of Shiraz University of Medical Sciences (Approval ID: IR.SUMS.SCHEANUT.REC.1403.044).

2.2. Data Gathering Tools

2.2.1. Demographic/Occupational Questionnaire

A questionnaire was used to gather demographic and occupational data, including details such as age (in years), height (in centimeters), weight (in kilograms), years of work experience, daily working hours, and marital status (single or married).

2.2.2. Persian Version of the Cornell Musculoskeletal Discomfort Questionnaire (P-CMDQ)

The Cornell questionnaire was initially developed by Hedge et al. in 1999 [25]. This questionnaire is designed to measure the frequency, severity, and interference with work ability related to the last working week. It evaluates 12 body regions. Additionally, this questionnaire has demonstrated validity and reliability in ergonomic assessments, with the psychometric properties of the Persian version evaluated by Afifeh-zadeh Kashani et al. [26]. We used a weighting system to better identify the most serious issues. The scoring method for the Frequency Score is as follows:

- Never = 0
- 1-2 Times/Week = 1.5
- 3-4 Times/Week = 3.5
- Every Day = 5
- Several Times Every Day = 10

These Frequency Scores are then multiplied by the Discomfort Score (ranging from 1 to 3) and the Interference Score (also ranging from 1 to 3). The final score ranges from 0 to 90.

2.2.3. The Lifting Fatigue Failure Tool (LiFFT)

This tool is used for risk assessment of manual material handling tasks. It is based on fatigue failure theory, which evaluates the cumulative damage to materials subjected to repeated stress. The LiFFT tool has been validated using two well-established epidemiological databases, demonstrating strong links with lower back disorders and back pain [24]. Its goal is to assess the accumulated load on the lower back during a workday. Using the LiFFT cumulative damage measure, it estimates the probability of a job being classified as high-risk, defined as having 12 or more injuries per 200,000 work hours [27]. To operate the LiFFT tool, three factors are needed for each lifting task: 1) the weight of the load, 2) the maximum horizontal distance from the hip joint to the load's center during lifting (measured with a tape measure), and 3) the number of repetitions of the task throughout the workday. For jobs involving multiple lifting tasks, the tool adds together the cumulative damage of each task to determine the overall risk.

2.2.4. The Shoulder Work Assessment Tool (SWAT)

The Shoulder Risk Assessment is designed to evaluate risks related to occupational tasks that involve stressful shoulder exertions. Based on fatigue failure theory, this tool estimates cumulative damage by analyzing shoulder moments and loading cycles. To use the tool, three pieces of information are required for each shoulder task: 1) the weight held or force exerted by the hands, 2) the maximum horizontal distance from the acromion (the flat bone at the top of the shoulder) to the center of the hand or load during the task (measured with a tape measure), and 3) the total number of repetitions performed throughout the workday. For tasks involving pushing forward or backward, the measuring tape should be held vertically. Load weight should be divided between the hands, either evenly or unevenly, as estimated by the analyst if one shoulder bears more load. When measuring lever arms for both shoulders, the maximum lever arm for each shoulder must be assessed, as it may occur at different times during the task. The tool can

analyze single-task jobs, multi-task jobs by summing cumulative damage, or highly variable jobs using a binning procedure to group tasks by shoulder moments. The output indicates the probability of shoulder symptoms severe enough to require medical attention [22].

2.2.5. The Distal Upper Extremity Tool (DUET)

The DUET tool assesses risks associated with tasks involving the distal upper extremities, based on fatigue failure theory. Research provides substantial evidence that upper extremity disorders and other musculoskeletal disorders result from cumulative damage due to repetitive stress [28]. The DUET tool has been validated using a cross-sectional epidemiological database, showing strong associations with upper extremity outcomes. The tool's primary goal is to determine the cumulative upper extremity load experienced during a workday, calculating the probability of symptoms severe enough to prompt a first-time medical visit [23].

To use the DUET tool, two pieces of information are required for each task: 1) the intensity rating of the exertion, and 2) the number of task repetitions throughout the workday. Exertion intensity can be assessed subjectively by the worker using the 10-point RPE (OMNI-RES) scale, were workers rate effort, strain, discomfort, or fatigue [29], or by observers using descriptors from the Strain Index [30]. The tool can analyze mono-task jobs, multi-task jobs by summing cumulative damage, or highly variable jobs using a binning procedure to group tasks by exertion level, providing actionable insights for task redesign.

2.3. Implementation of the Study

Participants completed an informed consent form, a demographic and occupational questionnaire, and the P-CMDQ after being briefed on the study process. Following this, assessors collected and recorded data using the LiFFT, SWAT, and DUET tools. The collected information was then used to calculate the cumulative loads on the lower back, shoulders, and distal upper extremities of the mechanics, utilizing these tools.

2.4. Statistical Analysis

The data were analyzed using version 16 of the Statistical Package for the Social Sciences (SPSS) software. Descriptive statistics were calculated for the variables of interest. To evaluate the normality of the data, the Kolmogorov-Smirnov and Shapiro-Wilk tests were conducted, and the results indicated significant deviations from normality. Partial correlation analyses, adjusted for age and Body Mass Index (BMI), were employed to examine relationships between risk levels for the lower back, shoulders, and distal upper extremities, assessed with the LiFFT, SWAT, and DUET tools, and musculoskeletal discomfort reported through the P-CMDQ. A significance level of 5% ($\alpha = 0.05$) was used for all statistical analyses.

3. Results

Table 1 provides an overview of the personal and occupational characteristics of the automotive mechanics participating in the study.

The frequency of WMSDs reported in the 12 months before the study is detailed below: 'every day' (14.6%), 'several times per week' (10.2%),

Table 1. Some personal and occupational details of the participants (n=157).

Quantitative variable	Mean ± SD†	Min–Max
Age (years)	34.03±8.64	20-60
Weight (kg)	79.05±10.63	59-120
Height (cm)	175.17±6.82	159-190
BMI* (kg.m^2)	25.75±2.99	18.21-34.89
Job experience (years)	11.58±9.12	1-30
Working hours per day	11.18±1.51	6-12
Qualitative variable	**No. (%)**	
Marital status		
Single	64 (40.8)	
Married	93 (59.2)	
Education level		
High school diploma or less	126 (80.3)	
Post-secondary education	31 (19.7)	

**Body Mass Index.*
† Standard Deviation.

'several times per month' (19.1%), 'several times per year' (18.5%), 'several times every few years' (12.7%), and 'only once' (24.8%) of participants. The duration of WMSDs during the 12 months leading up to the study is outlined as follows: '0 days' (26.1%), '7 days' (37%), '8-30 days' (16.6%), 'more than 30 days but not every day' (7%), and 'every day' (13%).

Table 2 shows the reported musculoskeletal discomfort in various body regions of automotive mechanics over the past week, as assessed by the P-CMDQ. The highest discomfort is found in the lower back, shoulders, and hands regions among the participants.

Table 3 presents the results obtained from the LiFFT, SWAT, and DUET tools, which analyzed the lower back, shoulders, and distal upper extremities. For a more detailed assessment, the results are categorized into three ranges: 0–33% (low), 34–66% (moderate), and 67–100% (high). As shown, the cumulative loads fall within the high range for the lower back in 42.7% of cases, the shoulders in 40.8% of cases, and the distal upper extremities in 36.3% of cases.

Table 4 displays the Partial correlations between risk levels for the lower back, shoulders, and distal upper extremities, assessed using the LiFFT, SWAT, and DUET tools, and musculoskeletal discomfort reported through the P-CMDQ for 157 participants. A strong correlation was observed between cumulative load on the lower back and perceived discomfort in this region (r = 0.730), whereas the correlations for the shoulders (r = 0.611) and distal upper extremities (r = 0.537) were moderate. These findings suggest that elevated risk levels, as identified by the assessment tools, correspond to higher discomfort scores in the respective regions of the P-CMDQ [31].

Table 2. Reported musculoskeletal discomfort according to the P-CMDQ (n=157).

Body region	Mean ± SD	Min–Max
Neck	34.3 ± 7.0	0-90
Shoulder	59.2 ± 8.7	0-90
Upper back	30.7 ± 8.1	0-90
Arm	27.7 ± 7.1	0-90
Lower back	66.0 ± 11.2	0-90
Forearm	23.1 ± 6.8	0-90
Hand	46.1 ± 7.0	0-90
Hip	25.3 ± 3.2	0-90
Thigh	23.7 ± 9.9	0-90
Knee	38.0 ± 9.7	0-90
Shank	20.3 ± 6.0	0-90
Foot	38.1 ± 7.0	0-90

Table 3. Risk levels for the lower back, shoulders, and distal upper extremities were assessed using the LiFFT, SWAT, and DUET tools (n=157).

	Low	Moderate	High
	No. (%)	No. (%)	No. (%)
Lower back	46 (29.3)	44 (28)	67 (42.7)
Shoulders	54 (34.4)	39 (24.8)	64 (40.8)
Distal upper extremities	56 (35.7)	44 (28)	57 (36.3)

Table 4. Partial correlations between risk levels for the lower back, shoulders, and distal upper extremities, assessed by the LiFFT, SWAT, and DUET tools, and musculoskeletal discomfort reported via the P-CMDQ (n=157).

	Discomfort in lower back	
	r	p-value*
Cumulative load on the lower back	0.730	<0.001
	Discomfort in shoulders	
	r	p-value*
Cumulative load on the shoulders	0.611	<0.001
	Discomfort in distal upper extremities	
	r	p-value*
Cumulative load on the distal upper extremities	0.537	<0.001

**Partial correlation analyses, adjusted for age and Body Mass Index (BMI).*

4. Discussion

This study aimed to assess cumulative loads affecting the lower back, shoulders, and distal upper extremities among automotive mechanics. Musculoskeletal discomfort, reported through the

P-CMDQ, showed significant issues in these areas, with the lower back being the most affected, followed by the shoulders and hands. Risk assessments using the LiFFT, SWAT, and DUET tools indicated considerable occupational hazards. Specifically, 42.7% of participants had high-risk profiles for lower back disorders, suggesting their jobs are likely to cause 12 or more lower back injuries per 200,000 hours worked, as evaluated by the LiFFT tool [24,27], which has shown strong links to manual materials handling (MMH) and the Quick Exposure Check (QEC) technique [32]. Likewise, 40.8% of participants displayed elevated risk levels for shoulder disorders, implying a significant chance of developing shoulder symptoms severe enough to need medical attention, as calculated by the SWAT [22]. Additionally, 36.3% of participants exhibited high-risk profiles for distal upper extremity disorders, indicating a notable likelihood of symptoms in these areas requiring medical consultation, as assessed by the DUET tool [23]. Partial correlations, ranging from 0.537 to 0.730, were seen between the risk levels for all three regions and the corresponding P-CMDQ discomfort scores, confirming the effectiveness of LiFFT, SWAT, and DUET in identifying manual handling hazards [31]. These results underscore the importance of targeted ergonomic measures to reduce risks linked to these high-hazard tasks.

The high rate of WMSDs among automotive mechanics can be linked to interconnected factors. The physical requirements of the job, including repetitive tasks, awkward postures, and heavy lifting, accumulate loads that raise the risk of WMSDs in the lower back, shoulders, and distal upper extremities. During vehicle maintenance, mechanics often adopt non-neutral postures, such as bending or twisting while working underneath or beside a vehicle, which contribute to low back pain [13]. Personal traits, such as higher BMI and more years of job experience, also increase vulnerability to WMSDs [17]. The occurrence of WMSDs varies across different job roles, with risk rising when multiple factors occur simultaneously [12]. Importantly, vehicle repair workers without professional training are twice as likely to develop WMSDs compared to trained workers, highlighting the importance of proper training in injury prevention [33]. Figure 1 shows the main task categories, related physical demands, and their musculoskeletal impacts in automotive mechanics.

Recent studies have reported a high prevalence of WMSDs among vehicle repair workers, with an overall rate of 47.7% (95% CI, 42.7–53.2%). Lower

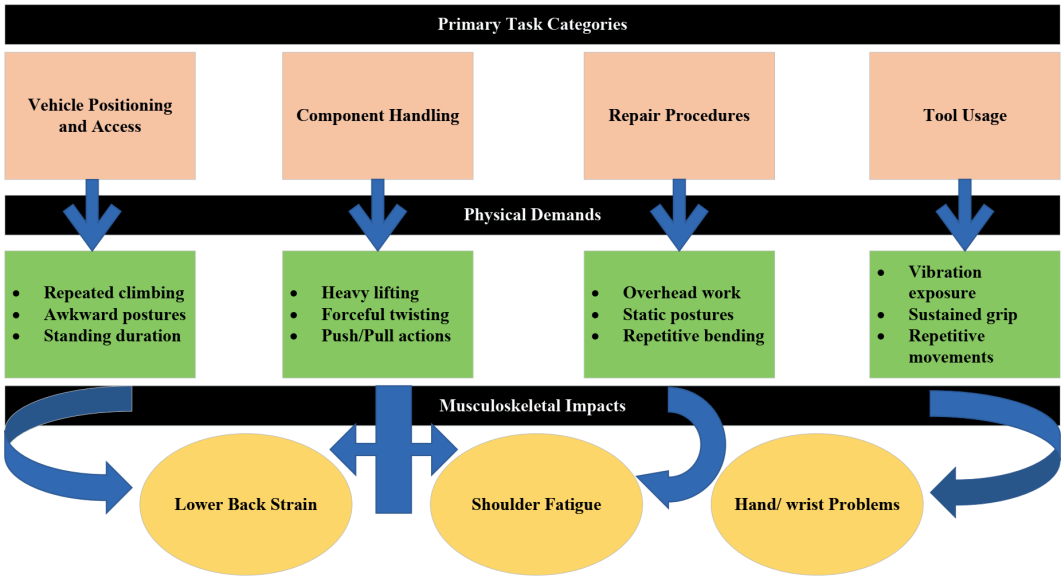

Figure 1: Primary task categories associated physical demands, and their musculoskeletal impacts in automotive mechanics.

back pain is the most common, affecting 62.8% of workers, followed by shoulder pain at 61% [18]. Abaraogu et al. reported a 76.02% prevalence of back pain, with 63.3% of individuals experiencing activity limitations [13]. He et al.'s meta-analysis confirmed that the lower back (36.5%), shoulders (31.4%), and wrist/hand (26.6%) are the most affected regions [17]. Zhang et al. reported similar patterns, with lower back (17.1%) and shoulder (14.5%) issues being prevalent [12]. WMSDs in these areas are linked to work absences [34]. Hernandez et al. found a mean Rapid Entire Body Assessment (REBA) technique score of 10.49 among truck mechanics, indicating very high risk [16]. Our study of 157 automotive mechanics confirms these findings, showing high-risk profiles for the lower back, shoulders, and distal upper extremities, emphasizing the need for ergonomic interventions.

4.1. Strengths and Limitations

The strengths of this study lie in its use of the LiFFT, SWAT, and DUET tools, which are specifically designed to quantify cumulative loads and risk levels of WMSDs in the lower back, shoulders, and distal upper extremities, respectively. These tools provide accurate risk assessments by incorporating specific inputs, such as load weight, distance, and task frequency for LiFFT and SWAT, and exertion intensity for DUET. Additionally, correlations between tool-assessed risk levels and P-CMDQ discomfort scores offer valuable insights into occupational hazards among automotive mechanics.

This study has several limitations. First, the sample only included male mechanics (n=157), which limits how well the results can be applied to female or mixed-gender groups and may overlook gender-specific risk factors. Second, the sample size might be too small to detect subtle trends in correlation analyses, and a larger sample could improve statistical power. Third, the subjective nature of the assessments—especially the reliance on self-reported data from tools like the P-CMDQ and DUET—may be prone to bias. Adding objective biomechanical measures in future research could improve validity. Finally, the study could not measure hand-arm vibration (HAV) exposure because of equipment limitations, which restricted the analysis of this potential WMSD risk factor.

4.2. Practical Strategies for Reducing WMSDs in Automotive Mechanics

The following recommendations aim to enhance workplace ergonomics and reduce musculoskeletal disorders among automotive mechanics. Ergonomic awareness and training: Mechanics should participate in training programs about ergonomics. Learning how to maintain proper body alignment, use tools effectively, and adopt safe lifting techniques can significantly decrease physical strain during daily tasks [33].

Use ergonomic tools: Mechanics should consider using specialized ergonomic tools and equipment designed to lessen physical effort. Tools that require less force, adjustable work surfaces, and lifting aids can help reduce strain on the back and shoulders.

Workstation adjustments: Employers should assess and improve workstations to make them ergonomically friendly. This includes adjusting the height and layout of work areas to reduce awkward postures and repetitive motions. Incorporate regular breaks: Mechanics should be encouraged to take short, frequent breaks during their shifts. These breaks allow for stretching and repositioning, helping to relieve muscle tension and prevent fatigue.

Strengthening and Flexibility Exercises: Establishing a routine that incorporates targeted strengthening and stretching exercises can greatly benefit mechanics by enhancing physical resilience, flexibility, and reducing musculoskeletal discomfort. Recent studies indicate that structured workplace stretching programs—especially when integrated into daily routines, such as during mid-shift or break times—can decrease fatigue and strain, improving worker well-being and performance [35, 36].

Monitoring Health: Regular health assessments focusing on musculoskeletal conditions can help identify issues early. Proactive monitoring enables prompt intervention, helping to prevent further problems and support overall well-being. Open Communication about Symptoms: Creating a culture where mechanics feel comfortable reporting

discomfort or MSD symptoms without fear of repercussions is essential. Early reporting enables timely interventions and adjustments to work.

Prioritize Recovery and Rest: Mechanics should be encouraged to focus on proper rest and recovery, especially after demanding shifts involving heavy lifting or awkward postures. Adequate recovery time is vital for muscle repair and long-term health.

Job Rotation Opportunities: Implementing job rotation strategies within a comprehensive ergonomic program can help distribute physical demands and reduce localized musculoskeletal stress among mechanics. However, recent evidence suggests that job rotation alone might not be sufficient to lower musculoskeletal disorders, particularly when high-risk tasks are involved. In such cases, redesigning and improving high-risk tasks should be prioritized. Once overall risk levels decrease, job rotation can then more effectively help reduce physical overload and support worker health [37, 38]. Management Support: Management must recognize the risks associated with MSDs and actively support efforts to mitigate these risks. This includes investing in ergonomic solutions and emphasizing health and safety in workplace policies.

4.3. Recommendations for Future Studies

Future studies should involve larger and more diverse populations to improve the generalizability of the findings. It is recommended that upcoming research compare the sensitivity and validity of LiFTT, DUET, and SWAT across different workplace settings and task types. To reduce self-reporting bias and enhance measurement accuracy, using objective ergonomic assessment methods—such as motion capture systems and wearable sensors—is recommended. Additionally, future research could examine including hand-arm vibration (HAV) as a variable, analyzing its presence and potential role in musculoskeletal complaints. Ultimately, intervention-focused studies evaluating the effectiveness of ergonomic improvements and worker training programs could identify practical strategies to reduce the occurrence and severity of work-related musculoskeletal disorders in physically demanding jobs.

5. Conclusion

This study highlights a high prevalence of WMSDs among automotive mechanics, especially affecting the lower back, shoulders, and distal upper extremities. Using the P-CMDQ and standardized tools (LiFFT, SWAT, DUET), our results show that 42.7%, 40.8%, and 36.3% of participants face high risks for WMSDs in the lower back, shoulders, and distal upper extremities, respectively. These risks are mainly caused by workplace factors such as repetitive movements and awkward postures. The positive correlations between tool-assessed risks and P-CMDQ discomfort scores support the effectiveness of the tools in identifying manual handling hazards. These findings underscore the pressing need for ergonomic interventions and preventive measures to enhance the health and productivity of automotive mechanics.

Funding: This work was supported by the Shiraz University of Medical Sciences under Grant 30845.

Institutional Review Board Statement: The study was conducted according to the guidelines of the declaration of Helsinki and approved by the ethics committee of Shiraz University of Medical Sciences (Ethics Code: IR.SUMS.SCHEANUT.REC.1403.044).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Acknowledgments: The authors would like to thank automotive mechanics who participated in the study.

Declaration of Interests: The authors declare no conflict of interests.

Author Contribution Statement: H.D. was involved in the study design, interpretation of the data, drafting of the manuscript. R.S. and N.R. were involved in the study design, data collection, interpretation of the data, drafting of the manuscript. N.S. and F.N. were involved in the data collection, analysis and interpretation of the data and drafting of the manuscript. L.G. was involved in the analysis and interpretation of the data and drafting of the manuscript. All authors have read and approved the final manuscript.

Declaration of Generative AI: During the preparation of this work the authors used ChatGPT 4 in order to improve readability, language, and grammar of some of the

text in the submitted manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

1. Wadsworth E, Walters D. Safety and Health at the Heart of the Future of Work: Building on 100 Years of Experience. International Labour Organization; 2019 978-92-2-133152-0. https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@dgreports/@dcomm/documents/publication/wcms_686645.pdf
2. Industrial Hygiene. Occupational Safety and Health Administration; 1998. https://www.osha.gov/publications/OSHA3143
3. Carrillo-CastrilloJA,Pérez-MiraV,Pardo-FerreiraMdC, Rubio-Romero JC. Analysis of required investigations of work-related musculoskeletal disorders in Spain. *Int J Environ Res Public Health*. 2019;16(10):1682. Doi: 10.3390/ijerph16101682
4. Arghami S, Kalantari R, Ahmadi Kionani E, Zanjirani Farahani A, Kamrani M. Assessing prevalence of musculoskeletal disorders in women workers in an automobile manufacturing assembly line. *J Hum Environ Health Promot*. 2016;1(2):75-79. Doi: jhehp/article-1-28-en.html
5. Ribeiro T, Serranheira F, Loureiro H. Work related musculoskeletal disorders in primary health care nurses. *Appl Nurs Res.* 2017;33:72-77. Doi: 10.1016/j.apnr.2016.09.003
6. Luttmann A, Jager M, Griefahn B, Caffier G, Liebers F. Preventing musculoskeletal disorders in the workplace. Protecting Workers' Health Series, No 5. World Health Organization, Geneva, Switzerland, 2003.
7. Introduction to work related musculoskeletal disorders. European Agency for Safety and Health at Work.; 2007 https://osha.europa.eu/en/publications/factsheet-71-introduction-work-related-musculoskeletal-disorders
8. Piedrahita H. Costs of work-related musculoskeletal disorders (MSDs) in developing countries: Colombia case. *Int J Occup Saf Ergon*. 2006;12(4):379-386. Doi: 10.1080/10803548.2006.11076696
9. Long MH, Johnston V, Bogossian F. Work-related upper quadrant musculoskeletal disorders in midwives, nurses and physicians: A systematic review of risk factors and functional consequences. *Appl Ergon*. 2012;43(3):455-467. Doi: 10.1016/j.apergo.2011.07.002
10. Al-Hourani Z, Nazzal M, Khader Y, Almhdawi K, Bibars AR. Work-related musculoskeletal disorders among Jordanian dental technicians: Prevalence and associated factors. *Work*. 2017;56(4):617-623. Doi: 10.3233/WOR-172524
11. Chinedu OO, Henry AT, Nene JJ, Okwudili JD. Work-related musculoskeletal disorders among office workers in higher education institutions: A cross-sectional study. *Ethiop J Health Sci*. 2020;30(5):715-724. Doi: 10.4314/ejhs.v30i5.10
12. Zhang H, Deng H, Jia N, et al. Epidemiological study of work-related musculoskeletal disorders and related risk factors among automobile maintenance workers. *Work*. 2023;76(3):1219-1231. Doi: 10.3233/WOR-220412
13. Abaraogu UO, Ezema CI, Igwe S, Egwuonwu AV, Okafor UC. Work-related back discomfort and associated factors among automotive maintenance mechanics in Eastern Nigeria: a cross sectional study. *Work*. 2016;53(4):813-823. Doi: 10.3233/WOR-162247
14. Hafez KA. Occupational discomfort and injuries among automotive technicians in Jeddah, Saudi Arabia: a cross-sectional study. *Work*. 2022;73(4):1203-1216. Doi: 10.3233/WOR-210522
15. Hembecker PK, Reis DC, Konrath AC, Gontijo LA, Merino EA. Investigation of musculoskeletal symptoms in a manufacturing company in Brazil: a cross-sectional study. *Braz J Phys Ther*. 2017;21(3):175-183. Doi: 10.1016/j.bjpt.2017.03.014
16. Oliver Hernández C, Li S, Aguado Benedí MJ, Mateo Rodríguez I. New challenges regarding the intervention of musculoskeletal risk in truck service garages. *Sustainability*. 2021;14(1):181. Doi: 10.3390/su14010181
17. He X, Xiao B, Wu J, Chen C, Li W, Yan M. Prevalence of work-related musculoskeletal disorders among workers in the automobile manufacturing industry in China: a systematic review and meta-analysis. *BMC Public Health*. 2023;23(1):2042. Doi: 10.1186/s12889-023-16896-x
18. Tamene A, Mulugeta H, Ashenafi T, Thygerson SM. Musculoskeletal disorders and associated factors among vehicle repair workers in Hawassa City, Southern Ethiopia. *J Environ Public Health*. 2020;2020(1):9472357. Doi: 10.1155/2020/9472357
19. Patel B, Patel M, Parmar A, Rathor D. Prevalence of Work-Related Musculoskeletal Pain among Garage Workers of Bharuch District: A Cross-Sectional Survey. *Int J Health Sci Res*. 2023. Doi: 10.52403/ijhsr.20230735
20. Coenen P, Kingma I, Boot CR, Bongers PM, van Dieën JH. Cumulative mechanical low-back load at work is a determinant of low-back pain. *Occup Environ Med*. 2014;71(5):332-337. Doi: 10.1136/oemed-2013-101862
21. Johnen L, Schaub M, Mertens A, Nitsch V, Brandl C. Can cumulative loading estimates be used to assess the collective occupational risk of MSD? Evaluation of calculation methods for spinal cumulative loading. *Int J Ind Ergon.* 2022;92:103361. Doi: 10.1016/j.ergon.2022.103361
22. Bani Hani D, Huangfu R, Sesek R, Schall Jr MC, Davis GA, Gallagher S. Development and validation of a cumulative exposure shoulder risk assessment tool based on fatigue failure theory. *Ergonomics*. 2021;64(1):39-54. Doi: 10.1080/00140139.2020.1811399

23. Gallagher S, Schall Jr MC, Sesek RF, Huangfu R. Validation of A Fatigue Failure-based Risk Assessment Tool for Distal Upper Extremity MSDs. SAGE Publications Sage CA: Los Angeles, CA; 2017:911-913.
24. Gallagher S, Sesek RF, Schall Jr MC, Huangfu R. Development and validation of an easy-to-use risk assessment tool for cumulative low back loading: The Lifting Fatigue Failure Tool (LiFFT). *Appl Ergon*. 2017;63:142-150. Doi: 10.1016/j.apergo.2017.04.016
25. Hedge A, Morimoto S, Mccrobie D. Effects of keyboard tray geometry on upper body posture and comfort. *Ergonomics*. 1999;42(10):1333-1349. Doi: 10.1080/00140 1399184983
26. Afifehzadeh-Kashani H, Choobineh A, Bakand S, Gohari M, Abbastabar H, Moshtaghi P. Validity and reliability of farsi version of Cornell Musculoskeletal Discomfort Questionnaire (CMDQ). *Iran Occup Health*. 2011;7(4): 69-75.
27. Marras WS, Lavender SA, Leurgans SE, et al. The role of dynamic three-dimensional trunk motion in occupationally related. *Spine*. 1993;18(5):617-628. Doi: 10.1097 /00007632-199304000-00015
28. Gallagher S, Schall MC Jr. Musculoskeletal disorders as a fatigue failure process: evidence, implications and research needs. *Ergonomics*. 2017;60(2):255-269. Doi: 10.1080 /00140139.2016.1208848.
29. Robertson RJ, Goss FL, Rutkowski J, et al. Concurrent validation of the OMNI perceived exertion scale for resistance exercise. *Med Sci Sports Exer.* 2003;35(2):333-341. Doi: 10.1249/01.mss.0000048831.15016.2a
30. Steven Moore J, Garg A. The strain index: a proposed method to analyze jobs for risk of distal upper extremity disorders. *Am Ind Hyg Ass J*. 1995;56(5):443-458. Doi: 10.1080/15428119591016863
31. Chiorri C, Garbarino S, Bracco F, Magnavita N. Personality traits moderate the effect of workload sources on perceived workload in flying column police officers. *Frontiers in psychology*. 2015;6:1835. Doi: 10.3389/fpsyg.2015 .01835
32. Eynipour A, Arjmand N, Dianat I, Soltanian AR, Heidarimoghadam R. Assessing Musculoskeletal Disorder Risks in an Automobile Part Manufacturing Factory: A Comparison Study of Biomechanical and Ergonomic Tools. *Health Scope.* 2024;13(2). Doi: 10.5812 /healthscope-139610
33. Boini S, Colin R, Grzebyk M. Effect of occupational safety and health education received during schooling on the incidence of workplace injuries in the first 2 years of occupational life: a prospective study. *BMJ Open*. 2017;7(7):e015100. Doi: 10.1136/bmjopen-2016 -015100
34. Torp S, Riise T, Moen B. Work-related musculoskeletal symptoms among car mechanics: a descriptive study. *Occup Med.* 1996;46(6):407-413. Doi: 10.1093 /occmed/46.6.407
35. Alqhtani R, Ahmed H, Alshahrani A, Khan A, Khan A. Effects of whole-body stretching exercise during lunch break for reducing musculoskeletal pain and physical exertion among healthcare professionals. *Medicina*. 2023;59(5):910. Doi: 10.3390/medicina59050910
36. Gasibat Q, Rani B, Causevic D, et al. Impact of stretching exercises on work-related musculoskeletal disorders: A systematic review. *Int J Kinesiol Sport Sci.* 2023;11(3):8-22. Doi: 10.7575/aiac.ijkss.v.11n.3p.8
37. Mehdizadeh A, Vinel A, Hu Q, Schall Jr M, Gallagher S, Sesek R. Job rotation and work-related musculoskeletal disorders: a fatigue-failure perspective. *Ergonomics*. 2020;63(4):461-476. Doi: 10.1080/0014 0139.2020.1717644
38. Comper M, Padula R. The effectiveness of job rotation to prevent work-related musculoskeletal disorders: protocol of a cluster randomized clinical trial. *BMC MusculoskeletDisord*.2014;15:1-6.Doi:10.1186/1471-2474-15-170